STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 99177

TO: Rebecca Prouty **Location: cm-1/10a13/10d01** **Art Unit: 1652** **Friday, August 01, 2003** **Case Serial Number: 10/089986**	**From: Toby Port** **Location: Biotech-Chem Library** **CM1-6A04** **Phone: 308-3534** **toby.port@uspto.gov**

Search Notes

Dear Examiner Prouty,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:29:07 ; Search time 268 Seconds
(without alignments)
7112.734 Million cell updates/sec

Title: US-10-089-986-1
Perfect score: 924
Sequence: 1 atggggccagttgttgaaaa.....cgggagttaaagtgaagcta 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq3:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		74.4	8.1	1356	11	US-09-880-107-1648	Sequence 1648, Ap
2		44	4.8	421	11	US-09-960-352-6109	Sequence 6109, Ap
3		43.6	4.7	471	12	US-09-918-995-15653	Sequence 15653, A
4		39	4.2	377	11	US-09-960-352-8453	Sequence 8453, Ap
5	c	38.6	4.2	11812	15	US-10-239-676-209	Sequence 209, App
6		37.2	4.0	3290	15	US-10-027-632-116285	Sequence 116285,
7		37.2	4.0	3290	15	US-10-027-632-116286	Sequence 116286,
8		35.8	3.9	1114	15	US-10-099-278-85	Sequence 85, Appl
9		35.8	3.9	1241	15	US-10-099-278-42	Sequence 42, Appl
10		35.8	3.9	2132	15	US-10-099-278-39	Sequence 39, Appl
11		35.2	3.8	32038	10	US-09-764-878-292	Sequence 292, App
12		35.2	3.8	32038	15	US-10-079-854-292	Sequence 292, App
13		35	3.8	520	14	US-10-146-731-144	Sequence 144, App
14		35	3.8	520	15	US-10-123-155-144	Sequence 144, App
15		34.8	3.8	630	11	US-09-974-300-2378	Sequence 2378, Ap
16		34.8	3.8	162450	15	US-10-071-179-1	Sequence 1, Appli
17		34.6	3.7	1560	12	US-09-995-898A-28	Sequence 28, Appl
18	c	34.4	3.7	637	15	US-10-027-632-233098	Sequence 233098,
19		34.4	3.7	4923	8	US-08-781-986A-551	Sequence 551, App
20	c	34.4	3.7	640681	11	US-09-790-988-1	Sequence 1, Appli
21	c	33.8	3.7	586	15	US-10-027-632-259845	Sequence 259845,
22	c	33.8	3.7	1321	11	US-09-938-842A-4627	Sequence 4627, Ap
23	c	33.8	3.7	2003	11	US-09-887-576-24	Sequence 24, Appl
24		33.6	3.6	662	15	US-10-027-632-137361	Sequence 137361,
25	c	33.6	3.6	43360	15	US-10-114-170-206	Sequence 206, App
26	c	33.6	3.6	45325	15	US-10-114-170-261	Sequence 261, App
27		33.4	3.6	663	15	US-10-027-632-33841	Sequence 33841, A
28		33.4	3.6	663	15	US-10-027-632-33842	Sequence 33842, A
29	c	33.2	3.6	491	11	US-09-796-692-6502	Sequence 6502, Ap
30	c	33.2	3.6	491	15	US-10-040-862-6502	Sequence 6502, Ap
31		33.2	3.6	1004	15	US-10-027-632-71236	Sequence 71236, A
32		33.2	3.6	1004	15	US-10-027-632-71237	Sequence 71237, A
33		33.2	3.6	1004	15	US-10-027-632-72284	Sequence 72284, A
34		33.2	3.6	2079	10	US-09-815-242-6598	Sequence 6598, Ap
35	c	33.2	3.6	402850	12	US-09-844-653-5	Sequence 5, Appli
36		33.2	3.6	640681	11	US-09-790-988-1	Sequence 1, Appli
37		33	3.6	454	15	US-10-027-632-142839	Sequence 142839,
38	c	33	3.6	397658	11	US-09-813-320-3	Sequence 3, Appli
39		32.8	3.5	966	12	US-09-764-891-8439	Sequence 8439, Ap
40		32.8	3.5	1087	11	US-09-938-842A-4021	Sequence 4021, Ap
41		32.8	3.5	1152	15	US-10-099-278-37	Sequence 37, Appl
42		32.8	3.5	2000	11	US-09-938-842A-2805	Sequence 2805, Ap
43	c	32.8	3.5	2171	11	US-09-331-631A-2	Sequence 2, Appli
44		32.8	3.5	2230	15	US-10-099-278-34	Sequence 34, Appl
45		32.8	3.5	2140405	15	US-10-027-632-76212	Sequence 76212, A

ALIGNMENTS

RESULT 1
US-09-880-107-1648
; Sequence 1648, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1648
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D31815
US-09-880-107-1648

Query Match 8.1%; Score 74.4; DB 11; Length 1356;
Best Local Similarity 48.6%; Pred. No. 9.1e-11;
Matches 297; Conservative 0; Mismatches 302; Indels 12; Gaps 3;

```
Qy      298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
            || ||     ||| ||||| || |  |||||||||| || |  ||||||   || |
Db      370 GTGGATAACGACAAGAAAAACAATCGCTTCAATGATGGGAAGGTGGATCCCGCCGGGAGG 429

Qy      358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCG---TAGGACCGGTCACTGGC 414
              |   | || || ||||||   ||  | | ||   | |   | | | | |   ||
Db      430 TACTTTGCTGGCACCATGGCTGAGGAAACAGCTCCAGCAGTTCTTGAGCGGCACCAGGGG 489
```

```
Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
Db        490 GCCCTGTACTCCCTCTTTCCTGATCACCACGTGAAAAAGTACTTTGACCAGGTGGACATT 549

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
Db        550 TCCAATGGTTTGGATTGGTCGCTAGACCACAAAATCTTCTATTACATTGACAGCCTGTCC 609

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
Db        610 TACTCCGTGGATGCCTTTGACTATGACCTGCAGACAGGACAGATCTCCAACCGCAGAAGT 669

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
Db        670 GTTTACAAGCTAGAAAAGGAAGAA------CAAATCCCAGATGGAATGTGTATTGATGCT 723

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
Db        724 GAGGGGAAGCTCTGGGTGGCCTGTTACAATGGAGGAAGAGTGATTCGTTTAGATCCTGTG 783

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
Db        784 ACAGGGAAA---AGACTTCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 840

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
Db        841 TTTGGAGGGAAGAATTACTCTGAAATGTATGTGACCTGCGCCCGGGATGGGATGGACCCC 900

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
Db        901 GAGGGTCTTTTGAGGCAACCTGAAGCTGGTGGAATTTTCAAGATAACTGGTCTGGGGGTC 960

Qy        895 AAAGGTTTCGC 905
Db        961 AAAGGAATTGC 971
```

RESULT 2
US-09-960-352-6109
; Sequence 6109, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6109
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB34-038-Q1-E1-G9
US-09-960-352-6109

Query Match 4.8%; Score 44; DB 11; Length 421;
Best Local Similarity 50.2%; Pred. No. 0.025;
Matches 135; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

```
Qy        631 CCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAG 690
Db        116 CCAGATGGAATGTGTATTGATGTTGAGGGGAAGCTCTGGGTGGCCTGTTACAATGGAGGA 175

Qy        691 CGAATTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCA 750
Db        176 AGAGTGATCCGTTTGGATCCTGAGACAGGGAAA---AGACTCCAAACTGTGAAGTTGCCT 232

Qy        751 GATCCTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACA 810
Db        233 GTTGATAAAACAACCTCATGCTGCTTCGGAGGGAAGGATTACTCTGAAATGTACGTGACC 292

Qy        811 TCTGCTGGTCTTCAGCTTGACGACAGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTC 870
Db        293 TGTGCCCGGGATGGGTTGGACTCCAAGGGTCTTCTGCAGCAACCTGAGGCTGGTGGAATT 352

Qy        871 TACAGAGTAACAGGTTTAGGCGTCAAAGG 899
Db        353 TTCAAGATAACTGGCCTAAGAGTCAAAGG 381
```

RESULT 3
US-09-918-995-15653
; Sequence 15653, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 15653
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(471)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-15653

Query Match 4.7%; Score 43.6; DB 12; Length 471;
Best Local Similarity 48.9%; Pred. No. 0.035;
Matches 180; Conservative 0; Mismatches 179; Indels 9; Gaps 2;

```
Qy        434 CTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGA 493
Db        112 CTGATCACCACGTGAAAAAGTACTTTGACCAGGTGGACATTTCCAATGGTTTGGATTGGT 171

Qy        494 GTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACG 553
Db        172 CGCTAGACCACAAAATCTTCTATTACATTGACAGCCTGTCCTACTCCGTGGATGCCTTTG 231

Qy        554 ATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGC 613
Db        232 ACTATGACCTGCAGACAGGACAGATCTCCAACCGCAGAAGTGTTTACAAGCTAGAAAAGG 291

Qy        614 ATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTG 673
Db        292 AAGAA------CAAATCCCAGATGGAATGTGTATTGATGCTGAGGGGAAGCTCTGGGTGG 345

Qy        674 CCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGG 733
Db        346 CCTGTTACAATGGAGGAAGAGTGATTCGTTTAGATCCTGTGACAGGGAAAAG---ACTTC 402

Qy        734 ATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGG 793
Db        403 AAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGCTTTGGAGGGAAGAATTACT 462

Qy        794 ATGAACTG 801
Db        463 CTGAAATG 470
```

RESULT 4
US-09-960-352-8453
; Sequence 8453, Application US/09960352
; Patent No. US20020137139A1

```
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8453
;   LENGTH: 377
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 36-LIB34-030-Q1-E1-A8
US-09-960-352-8453

  Query Match             4.2%;  Score 39;  DB 11;  Length 377;
  Best Local Similarity   51.1%;  Pred. No. 0.65;
  Matches  118;  Conservative    0;  Mismatches  110;  Indels    3;  Gaps    1;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||   | | |  ||||| ||  || |||||||  || || |  ||||||  ||| |
Db        145 GTAGATAAAGAGAAGAAAAACAATCGATTCAATGACGGGAAGGTGGATCCCGCTGGGAGA 204

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414
                |   | |||||  ||||||    ||  | | ||   | ||     |  |     |   ||
Db        205 TACTTTGCTGGTACCATGGCTGAGGAAACAGCTCCGGCAGTTTTGGAGCGCCGCCAAGGG 264

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
                | ||     |    ||||  |  | ||  ||| | ||      | |  | |  ||
Db        265 TCCCTGTACTCGCTCTTTCCTGACCACCATGTGGAAAAGTACTTTGACCAGGTGGACATC 324

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGAT 525
                | |||||| | |  |||  ||  ||    || |   | |||||  ||  ||
Db        325 TCCAATGGGTTGGGTTGGTCTATGGACCACAAAATCTTCTATTACATAAAT 375


RESULT 5
US-10-239-676-209/c
; Sequence 209, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
;  FILE REFERENCE: 5013.1003
;  CURRENT APPLICATION NUMBER: US/10/239,676
;  CURRENT FILING DATE:  2002-09-24
;  PRIOR APPLICATION NUMBER: PCT/EP01/03968
;     DE 10019058.8
;     DE 10019173.8
;     DE 10032529.7
;     DE 10043826.1
;  PRIOR FILING DATE: 2001-04-06
;     2000-04-06
;     2000-04-07
;     2000-06-30
;     2000-09-01
;  NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 209
;   LENGTH: 11812
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-209

  Query Match             4.2%;  Score 38.6;  DB 15;  Length 11812;
  Best Local Similarity   58.1%;  Pred. No. 5.2;
  Matches   68;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy         67 GATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCATAAATAT 126
              ||  ||  ||   | |||  |||| |  |  ||||   | | | |  |||  ||| ||||
Db       3100 GAAAATCTAAACAATACCCAATATATAATTAACACTCAATAAATATTTTTAAATATATAT 3041

Qy        127 GTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTCATT 183
               ||  |    | |||||||| || |||   |||  | |||||  |   ||  || |||
Db       3040 ATAAATAAAAAAAAAAAATAAACCTTTAAAAAATAATATAAAATTTATTTATTTATT 2984


RESULT 6
US-10-027-632-116285
; Sequence 116285, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 116285
;   LENGTH: 3290
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-116285

  Query Match             4.0%;  Score 37.2;  DB 15;  Length 3290;
  Best Local Similarity   59.4%;  Pred. No. 6.7;
  Matches   63;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy         51 AGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAA 110
              ||||||| | || |   | ||| ||     ||  | || ||| |  | |||  | | ||
Db       3060 AGAAGGTTCACAGTAATAACATCAAGAATTGAGGTGATGTTTAGAAGTCACAATTGTGAG 3119

Qy        111 AACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGT 156
              | || |  | | | ||||| ||| || |  |||| |||    ||||
Db       3120 AGCTCTAGAAATAGATGTATCTTTTCTGTCAAAACACATTAATTGT 3165


RESULT 7
US-10-027-632-116286
; Sequence 116286, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
```

```
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 116286
;   LENGTH: 3290
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-116286

  Query Match             4.0%;  Score 37.2;  DB 15;  Length 3290;
  Best Local Similarity   59.4%;  Pred. No. 6.7;
  Matches   63;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy         51 AGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAA 110
              ||||||| | || |   | ||| ||     ||  | || ||| |  | |||  | | ||
Db       3060 AGAAGGTTCACAGTAATAACATCAAGAATTGAGGTGATGTTTAGAAGTCACAATTGTGAG 3119

Qy        111 AACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGT 156
              | || |  | | | ||||| ||| || |  |||| |||    ||||
Db       3120 AGCTCTAGAAATAGATGTATCTTTTCTGTCAAAACACATTAATTGT 3165


RESULT 8
US-10-099-278-85
; Sequence 85, Application US/10099278
; Publication No. US20030106093A1
; GENERAL INFORMATION:
;  APPLICANT: Narva, Kenneth E.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Knuth, Mark
;  APPLICANT:  Pollard, Michael R.
;  APPLICANT:  Cardineau, Guy A.
;  APPLICANT:  Schwab, George E.
;  APPLICANT:  Michaels, Tracy E.
;  APPLICANT:  Finstad Lee, Stacey
;  APPLICANT:  Burmeister, Paula
;  APPLICANT:  Dojillo, Joanna
;  TITLE OF INVENTION: Pesticidal Proteins
;  FILE REFERENCE: MA-703C2D1
;  CURRENT APPLICATION NUMBER: US/10/099,278
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: US 09/378,088
;  PRIOR FILING DATE: 1999-08-20
;  PRIOR APPLICATION NUMBER: US 08/633,993
;  PRIOR FILING DATE: 1996-04-19
;  PRIOR APPLICATION NUMBER: US 08/844,188
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 130
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
;   LENGTH: 1114
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-10-099-278-85

  Query Match             3.9%;  Score 35.8;  DB 15;  Length 1114;
  Best Local Similarity   45.5%;  Pred. No. 9.5;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||   ||  |  |  |   |||  ||| || |    | || |   ||    ||   ||
Db        618 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 677

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
               |||  |||  |  ||| | ||| | |  || ||   |  |          |    | |
Db        678 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 737

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
                |  ||     | ||||  |    ||  | | |  ||||| |  |  |||||   |||
Db        738 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 797

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |      | |   | |||    | |    |   || |
Db        798 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 857

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725
              |     ||| ||  |||    |    ||||| | ||| |
Db        858 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 896


RESULT 9
US-10-099-278-42
; Sequence 42, Application US/10099278
; Publication No. US20030106093A1
; GENERAL INFORMATION:
;  APPLICANT: Narva, Kenneth E.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Knuth, Mark
;  APPLICANT:  Pollard, Michael R.
;  APPLICANT:  Cardineau, Guy A.
;  APPLICANT:  Schwab, George E.
;  APPLICANT:  Michaels, Tracy E.
;  APPLICANT:  Finstad Lee, Stacey
;  APPLICANT:  Burmeister, Paula
;  APPLICANT:  Dojillo, Joanna
;  TITLE OF INVENTION: Pesticidal Proteins
;  FILE REFERENCE: MA-703C2D1
;  CURRENT APPLICATION NUMBER: US/10/099,278
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: US 09/378,088
;  PRIOR FILING DATE: 1999-08-20
;  PRIOR APPLICATION NUMBER: US 08/633,993
;  PRIOR FILING DATE: 1996-04-19
;  PRIOR APPLICATION NUMBER: US 08/844,188
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 130
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
;   LENGTH: 1241
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: Undetermined nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (53)..(53)
;   OTHER INFORMATION: Undetermined nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (61)..(61)
;   OTHER INFORMATION: Undetermined nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (68)..(68)
;   OTHER INFORMATION: Undetermined nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (73)..(73)
;   OTHER INFORMATION: Undetermined nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (81)..(81)
;   OTHER INFORMATION: Undetermined nucleotide.
US-10-099-278-42

  Query Match             3.9%;  Score 35.8;  DB 15;  Length 1241;
  Best Local Similarity   45.5%;  Pred. No. 10;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;
```

```
Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||  ||  |  |  |  ||| ||| || |  | || |  ||  ||  ||
Db        705 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 764

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
              |||  |||  |  ||| | ||| | |  || ||  |  |         |     | |
Db        765 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 824

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
              |  ||     | ||||  |    ||  | | |  ||||| |  |  |||||  |||
Db        825 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 884

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |     | |   | |||    | |    |   || |
Db        885 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 944

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725
              |    ||| ||  |||    |    ||||| | ||| |
Db        945 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 983


RESULT 10
US-10-099-278-39
; Sequence 39, Application US/10099278
; Publication No. US20030106093A1
; GENERAL INFORMATION:
;  APPLICANT: Narva, Kenneth E.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Knuth, Mark
;  APPLICANT:  Pollard, Michael R.
;  APPLICANT:  Cardineau, Guy A.
;  APPLICANT:  Schwab, George E.
;  APPLICANT:  Michaels, Tracy E.
;  APPLICANT:  Finstad Lee, Stacey
;  APPLICANT:  Burmeister, Paula
;  APPLICANT:  Dojillo, Joanna
;  TITLE OF INVENTION: Pesticidal Proteins
;  FILE REFERENCE: MA-703C2D1
;  CURRENT APPLICATION NUMBER: US/10/099,278
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: US 09/378,088
;  PRIOR FILING DATE: 1999-08-20
;  PRIOR APPLICATION NUMBER: US 08/633,993
;  PRIOR FILING DATE: 1996-04-19
;  PRIOR APPLICATION NUMBER: US 08/844,188
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 130
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
;   LENGTH: 2132
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-10-099-278-39

  Query Match             3.9%;  Score 35.8;  DB 15;  Length 2132;
  Best Local Similarity   45.5%;  Pred. No. 13;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||  ||  |  |  |  ||| ||| || |  | || |  ||  ||  ||
Db       1401 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 1460

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
              |||  |||  |  ||| | ||| | |  || ||  |  |         |     | |
Db       1461 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 1520

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
              |  ||     | ||||  |    ||  | | |  ||||| |  |  |||||  |||
Db       1521 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 1580

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |     | |   | |||    | |    |   || |
Db       1581 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 1640

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725
              |    ||| ||  |||    |    ||||| | ||| |
Db       1641 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 1679


RESULT 11
US-09-764-878-292
; Sequence 292, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 292
;   LENGTH: 32038
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-878-292

  Query Match             3.8%;  Score 35.2;  DB 10;  Length 32038;
  Best Local Similarity   49.5%;  Pred. No. 84;
  Matches   91;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy        473 TAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGA 532
              || | ||||  |  |   ||   | | |||| |  |||||  ||  | |||  |
Db      17391 TAACTAATGACCATGGAGGGGAGATTAATTTAAGAAAAATTAATCTTTTTGTATTATTTC 17450

Qy        533 AAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGC 592
              || | |   | ||  |  ||  ||| | || |||||  | |   |||  |   | ||| ||
Db      17451 TAAAAGGTAAAAACTAAAACCTTTAAGGTGATTCTAAAGTTCACATGAAAGAGCAAAGGG 17510

Qy        593 CATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATG 652
              |       |  |  || | |  |||||      ||| |  |   | | || || |    |
Db      17511 CCAACAGCAGCTAAGAGATGTTTGAAGAATGAGGAAAAGCTATTTGCCATACTAGATTTA 17570

Qy        653 AGGA 656
              | ||
Db      17571 AAGA 17574


RESULT 12
US-10-079-854-292
; Sequence 292, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121C1
;  CURRENT APPLICATION NUMBER: US/10/079,854
;  CURRENT FILING DATE:  2002-02-22
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 292
;   LENGTH: 32038
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-079-854-292

  Query Match             3.8%;  Score 35.2;  DB 15;  Length 32038;
  Best Local Similarity   49.5%;  Pred. No. 84;
  Matches   91;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy        473 TAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGA 532
              || | ||||  |  |   ||   | | |||| |  |||||  ||  | |||  |
Db      17391 TAACTAATGACCATGGAGGGGAGATTAATTTAAGAAAAATTAATCTTTTTGTATTATTTC 17450
```

```
Qy        533 AAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGC 592
              || | |   | || |  ||  ||| | || ||||| | |   ||| |   | ||| ||
Db      17451 TAAAAGGTAAAAACTAAAACCTTTAAGGTGATTCTAAAGTTCACATGAAAGAGCAAAGGG 17510

Qy        593 CATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATG 652
              |     | |  || | |  |||||     ||| |  |   | | || || |    |
Db      17511 CCAACAGCAGCTAAGAGATGTTTGAAGAATGAGGAAAAGCTATTTGCCATACTAGATTTA 17570

Qy        653 AGGA 656
              | ||
Db      17571 AAGA 17574
```

RESULT 13
US-10-146-731-144
; Sequence 144, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 144
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-144

Query Match 3.8%; Score 35; DB 14; Length 520;
Best Local Similarity 10.4%; Pred. No. 11;
Matches 39; Conservative 111; Mismatches 224; Indels 0; Gaps 0;

```
Qy        379 ATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGAT 438
              |   :|:    :     :|: ::  :   ::    ::   : || :  |: : |:|:
Db         38 AFLEKYGYLNEQVPKAPTSTRFSDAIRAFQWVSQLPVSGVLDRATLRQMTRPRCGVTDTN 97

Qy        439 AAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAAT 498
                ||:     :   ::: ::::: :       : :     : :   :     ::| ::
Db         98 SYAAWAERISDLFARHRTKMRRKKRFAKQGNKWYKQHLSYRLVNWPEHLPEPAVRGAVRA 157

Qy        499 GATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTAT 558
                :: ::|   :    | |: : |   |:  : | | :|  ||  :    ::  |:
Db        158 AFQLWSNVSALEFWEAPATGPADIRLTFFQGDHNDGLGNAFDGPGGALAHAFLPRRGEAH 217

Qy        559 GATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAA 618
                : :  ::   :  :   :  |    :      : |  |    :::  |::|  :::::
Db        218 FDQDERWSLSRRRGRNLFVVLAHEIGHTLGLTHSPAPRALMAPYYKRLGRDALLSWDDVL 277

Qy        619 GTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTT 678
                : :|:   ::|:  |:   :   :::    |::      :: : : :   : :
Db        278 AVQSLYGKPLGGSVAVQLPGKLFTDFETWDSYSPQGRRPETQGPKYCHSSFDAITVDRQQ 337

Qy        679 TTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACC 738
```

```
                :   |:: : : |:  :   :     :          || ::  :|: :    |: ::
Db        338 QLYIFKGSHFWEVAADGNVSEPRPLQERWVGLPPNIEAAAVSLNDGDFYFFKGGRCWRFR 397

Qy        739 GTAAAAATACCAGA 752
              |   ::      :|
Db        398 GPKPVWGLPQLCRA 411
```

RESULT 14
US-10-123-155-144
; Sequence 144, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 144
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-144

Query Match 3.8%; Score 35; DB 15; Length 520;
Best Local Similarity 10.4%; Pred. No. 11;
Matches 39; Conservative 111; Mismatches 224; Indels 0; Gaps 0;

```
Qy        379 ATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGAT 438
              |   :|:    :     :|: ::  :   ::    ::   : || :  |: : |:|:
Db         38 AFLEKYGYLNEQVPKAPTSTRFSDAIRAFQWVSQLPVSGVLDRATLRQMTRPRCGVTDTN 97

Qy        439 AAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAAT 498
                ||:     :   ::: ::::: :       : :     : :   :     ::| ::
Db         98 SYAAWAERISDLFARHRTKMRRKKRFAKQGNKWYKQHLSYRLVNWPEHLPEPAVRGAVRA 157

Qy        499 GATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTAT 558
                :: ::|   :    | |: : |   |:  : | | :|  ||  :    ::  |:
Db        158 AFQLWSNVSALEFWEAPATGPADIRLTFFQGDHNDGLGNAFDGPGGALAHAFLPRRGEAH 217

Qy        559 GATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAA 618
                : :  ::   :  :   :  |    :      : |  |    :::  |::|  :::::
Db        218 FDQDERWSLSRRRGRNLFVVLAHEIGHTLGLTHSPAPRALMAPYYKRLGRDALLSWDDVL 277

Qy        619 GTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTT 678
                : :|:   ::|:  |:   :   :::    |::      :: : : :   : :
Db        278 AVQSLYGKPLGGSVAVQLPGKLFTDFETWDSYSPQGRRPETQGPKYCHSSFDAITVDRQQ 337

Qy        679 TTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACC 738
                :   |:: : : |:  :   :     :          || ::  :|: :    |: ::
Db        338 QLYIFKGSHFWEVAADGNVSEPRPLQERWVGLPPNIEAAAVSLNDGDFYFFKGGRCWRFR 397

Qy        739 GTAAAAATACCAGA 752
              |   ::      :|
```

```
Db        398 GPKPVWGLPQLCRA 411


RESULT 15
US-09-974-300-2378
; Sequence 2378, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2378
;   LENGTH: 630
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-2378

  Query Match             3.8%;  Score 34.8;  DB 11;  Length 630;
  Best Local Similarity   48.9%;  Pred. No. 14;
  Matches   93;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy        642 AACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAA 701
              |||  ||  ||  | ||| |||||. ||   ||| || ||||     |    || |  ||
Db        250 AACGCTTATTGGCGTGGGGAATTTGGGGACAGCCTTTCTCCATTACAATTTTATCAAAAA 309

Qy        702 AATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGT 761
              |  | |||  ||    |    | |||       |  | | ||| |   |   |  ||||
Db        310 CAATAATACAAAAATCGCCATGGCGTTTGACATCAATGAAGAAAAAATCGGAACCGAGGT 369

Qy        762 CACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCT 821
              |. |   ||  |  |    |  | ||||  |  ||||  |||| | ||   || || | |
Db        370 CGGCGGCGTTCCCGTCTATGACCTGAATAAGCTTGAAGAGCATATGACGGATGATGATAT 429

Qy        822 TCAGCTTGAC 831
              | | | | |
Db        430 CCCGGTCGCC 439


Search completed: August  1, 2003, 14:30:57
Job time : 279 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:14:17 ; Search time 64 Seconds
(without alignments)
641.269 Million cell updates/sec

Title: US-10-089-986-2
Perfect score: 1615
Sequence: 1 MGPVVEKIAELGKYTVGEGP.....HVYRVTGLGVKGFAGVKVKL 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1615	100.0	308	22	AAB82087	Firefly luciferin
2	889.5	55.1	309	23	ABB09720	Amino acid sequenc
3	859.5	53.2	307	23	ABB06339	Luciola lateralis
4	568	35.2	303	22	ABB63233	Drosophila melanog
5	550	34.1	303	22	ABB58550	Drosophila melanog
6	550	34.1	303	22	ABB58551	Drosophila melanog
7	550	34.1	303	22	ABB66692	Drosophila melanog
8	457.5	28.3	299	23	ABB83787	Mouse SMP30 SEQ ID
9	454.5	28.1	299	16	AAR75416	Rat regucalcin, a
10	448.5	27.8	299	16	AAR74219	Human ageing marke
11	448.5	27.8	299	18	AAW14475	Human senility mar
12	448.5	27.8	299	20	AAY06995	Senescence marker
13	238	14.7	293	21	AAB11651	A. vitis hypersens
14	238	14.7	340	21	AAB11650	A. vitis hypersens
15	195.5	12.1	107	23	ABP32045	Human RNA polymera
16	124.5	7.7	320	23	AAU76347	Z. mobilis aldonol
17	110	6.8	332	23	ABP39801	Staphylococcus epi
18	110	6.8	762	23	AAU78666	Human NOV5b protei
19	109	6.7	729	23	AAU78665	Human NOV5a protei
20	109	6.7	2009	22	ABB64069	Drosophila melanog
21	104	6.4	324	22	AAB99723	Staphylococcus aur
22	103.5	6.4	369	21	AAG07424	Arabidopsis thalia
23	103.5	6.4	371	21	AAG07423	Arabidopsis thalia
24	103.5	6.4	398	21	AAG07422	Arabidopsis thalia
25	97.5	6.0	553	21	AAY74581	Neisseria meningit
26	97.5	6.0	553	23	AAU72944	Neisseria meningit
27	97.5	6.0	695	23	AAU91286	Human NOV5e protei
28	97	6.0	371	21	AAG25423	Arabidopsis thalia
29	97	6.0	377	21	AAG25422	Arabidopsis thalia
30	96	5.9	419	22	ABB58148	Drosophila melanog
31	96	5.9	3117	21	AAY53667	Sequence gi/332818
32	95	5.9	265	22	AAG82348	S. epidermidis ope
33	94.5	5.9	553	21	AAY74580	Neisseria meningit
34	94	5.8	413	23	ABP40266	Staphylococcus epi
35	93	5.8	396	16	AAR76550	Thermostable alkal
36	93	5.8	449	23	ABP28423	Streptococcus poly
37	92.5	5.7	370	21	AAG06314	Arabidopsis thalia
38	92.5	5.7	385	21	AAG06313	Arabidopsis thalia
39	92	5.7	804	23	AAU91287	Human NOV5f protei
40	92	5.7	825	23	AAU91284	Human NOV5c protei
41	92	5.7	847	23	AAU91289	Human NOV5h protei
42	92	5.7	857	23	AAU91288	Human NOV5g protei
43	92	5.7	1012	23	AAU91285	Human NOV5d protei
44	92	5.7	4263	23	ABB97541	Novel human protei
45	92	5.7	4349	23	AAU79940	Human protocadheri

ALIGNMENTS

RESULT 1
AAB82087
ID AAB82087 standard; Protein; 308 AA.
XX
AC AAB82087;
XX
DT 26-JUN-2001 (first entry)
XX
DE Firefly luciferin regenerating protein.
XX
KW Firefly; luciferin regenerating protein; oxyluciferin; luciferase.
XX
OS Photinus pyralis.
XX
FH Key Location/Qualifiers
FT Misc-difference 281
FT /note= "Encoded by TTN"
XX
PN WO200125426-A1.
XX
PD 12-APR-2001.
XX
PF 22-SEP-2000; 2000WO-JP06527.
XX
PR 06-OCT-1999; 99JP-0285258.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Kurosawa K, Kajiyama N;
XX
DR WPI; 2001-266307/27.
DR N-PSDB; AAF86444.
XX

PT Luciferin regenerating protein and gene encoding it useful for
PT regenerating expensive luciferin from oxyluciferin and D-cysteine -
XX
PS Claim 7; Page 16-17; 21pp; Japanese.
XX
CC The present sequence is a firefly protein which regenerates luciferin by
CC using oxyluciferin and D-cysteine. This protein can be used for
CC regeneration of luciferin, a substrate for luciferase, used for ATP
CC (adenosine triphosphate) assays in both medical and food hygiene areas.
XX
SQ Sequence 308 AA;

Query Match 100.0%; Score 1615; DB 22; Length 308;
Best Local Similarity 100.0%; Pred. No. 8.4e-153;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNL 120

Qy        121 WTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 WTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRR 180

Qy        181 VDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQP 240

Qy        241 EVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG 300

Qy        301 FAGVKVKL 308
              ||||||||
Db        301 FAGVKVKL 308
```

RESULT 2
ABB09720
ID ABB09720 standard; Protein; 309 AA.
XX
AC ABB09720;
XX
DT 11-JUN-2002 (first entry)
XX
DE Amino acid sequence of a protein capable of regenerating luciferin.
XX
KW Luciferin regenerating protein; luciferen; oxyluciferin; luminescence;
KW ATP assay.
XX
OS Luciola cruciata.
XX
PN WO200210383-A1.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-JP06454.
XX
PR 28-JUL-2000; 2000JP-0228226.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Hirokawa K, Kurosawa K, Kajiyama N;
XX
DR WPI; 2002-180080/23.
DR N-PSDB; ABL41998.
XX
PT New Luciola cruciata-originated genes encoding proteins capable of
PT regenerating luciferin especially from oxyluciferin, for producing
PT recombinant DNAs and transformants -
XX
PS Claim 1; Page 18-19; 23pp; Japanese.
XX
CC The present sequence represents a protein capable of regenerating
CC luciferin. The protein can especially regenerate luciferen from
CC oxyluciferin. The protein can be added to the luciferin-luciferase
CC reaction system to sustain luminescence and reduce the amounts used.
CC The polynucleotide sequence is useful for producing recombinant
CC transformants, which are useful for the production of proteins useful
CC in assaying ATP in medical sciences and food hygiene.
XX
SQ Sequence 309 AA;

Query Match 55.1%; Score 889.5; DB 23; Length 309;
Best Local Similarity 56.6%; Pred. No. 2.8e-80;
Matches 176; Conservative 47; Mismatches 83; Indels 5; Gaps 4;

```
Qy          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60
              | | ||:| ||| | : | |||| |||:||||| | :: :||||: | :|  | ||  ||
Db          1 MAPTVEQIVELGTYLLAESPHWDDETQSLYFVDIVGRSVNKYVPTTKTHTQLKFDKNPSF 60

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIE--AIVNVEPHIKNNRLNDGKADPLG 118
              |||: |.  ||:|||||| :|||| |:||: ||  |: :  |    ||||||||||||
Db         61 IIPVKGCSDRFIVSLEREINLLTWDGASSAPSKIEKIAVFDNTPEKSENRLNDGKADPLG 120

Qy        119 NLWTGTMAIDAGLPVG-PVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSG 177
              ||| ||| : :     | || |:|  | ::|:|| | | : |:|||||| |||| |||||
Db        121 NLWAGTMNMGSDHTTGTPVRGTLSSL-SNKQVKEHVSEVCISNGLAWSKDLKKFYYIDSA 179

Qy        178 KRRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKIST 237
              |:||::|:||  ||:||::|||||:|| : | ||||||| ||||||| || :::|| |
Db        180 VRQVDQFDFDAKNLSLSNRQPLFTFDKHGIMGSPDGQTIDAEGNLWVATCQGDKVLKIDT 239

Qy        238 QQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLG 297
              || ||  |:||: |||||  ||  |: |:||:| ::|  :  | :  | :|:|||||
Db        240 STPETLLGIVEIPEHQVTSVCIGGAELNVLYVTTASIKLPGADETKPM-KGAIYKVTGLG 298

Qy        298 VKGFAGVKVKL 308
              |||  | :|||
Db        299 VKGLPGDRVKL 309
```

RESULT 3
ABB06339
ID ABB06339 standard; Protein; 307 AA.
XX
AC ABB06339;
XX
DT 29-MAY-2002 (first entry)
XX
DE Luciola lateralis luciferin regenerating protein SEQ ID NO:2.
XX
KW Luciola lateralis; luciferin; regeneration; Japanese firefly;
KW adenosine triphosphate; ATP; medical science; food hygiene;
KW luciferase; luminescence.
XX
OS Luciola lateralis.
XX
PN WO200210384-A1.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-JP06455.
XX
PR 28-JUL-2000; 2000JP-0228227.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Hirokawa K, Kurosawa K, Kajiyama N;
XX
DR WPI; 2002-227086/28.
DR N-PSDB; ABL49712.

XX
PT Luciola lateralis-originated genes encoding proteins capable of
PT regenerating luciferin especially from oxyluciferin, useful for
PT producing recombinant DNAs and transformants to give proteins useful in
PT assaying adenosine triphosphate -
XX
PS Claim 1; Page 18-19; 23pp; Japanese.
XX
CC The present sequence represents a protein capable of regenerating
CC luciferin, which is isolated from Luciola lateralis (Japanese firefly).
CC The gene encoding the protein capable of regenerating luciferin can be
CC used for producing recombinant DNAs and transformants, which can be
CC used for the production of proteins useful in assaying adenosine
CC triphosphate (ATP) in medical sciences and food hygiene. The protein
CC can be added to the luciferin-luciferase reaction system to sustain
CC luminescence and reduce their amounts used.
XX
SQ Sequence 307 AA;

Query Match 53.2%; Score 859.5; DB 23; Length 307;
Best Local Similarity 52.8%; Pred. No. 2.7e-77;
Matches 163; Conservative 58; Mismatches 85; Indels 3; Gaps 3;

```
Qy          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60
              | ||:|:'| |:  : :|||||||| |||:||||| :||: |||||| |::|   ::|  ||
Db          1 MSPVIEQITEVDNFQIGEGPHWDTETQSLYFVDILEKSIHKYVPSTKQHTKMILNKRPSF 60

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNL 120
              |||:  :  |||:||||:| :|||||||| |: :| || |:  |: |  ||||||  |||
Db         61 IIPIKETSDRFVISLERDICVLTWDGVSATPSHLETIVTVDTGIEGNTFNDGKADAFGNL 120

Qy        121 WTGTMAIDAGL-PVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKR 179
              | ||:     :   || ||:|| | ::|::: | ||| ::|||||: | || |:||| ||
Db        121 WAGTLYSKFDIEKQGPNTGTLYSL-SNKQLRKHISNIFLSNGLAWNKDSKKFYFIDSNKR 179

Qy        180 RVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQ 239
             :|::|||:  | ||| :|||| :|| : | || |||||  |||||: :| ::| | |:|
Db        180 TIDQFDYDSENLIISNCQPLFTLDKHGIQGLPDAQTIDENDNLWVAIVRGGKVINIGTKQ 239

Qy        240 PEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVK 299
              || ||  : :|:  :||| ||| ||||:||::|::  ::   | || | :|||||||||
Db        240 PESLLGVINMPESLITSVCFGGSKLDELYVTTSGIKEYETDSTK-LVKGGLYRVTGLGVK 298

Qy        300 GFAGVKVKL 308
              |    : |
Db        299 GLPAHRFSL 307
```

RESULT 4
ABB63233
ID ABB63233 standard; Protein; 303 AA.
XX
AC ABB63233;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 16491.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL07336.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 16491; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 303 AA;

Query Match 35.2%; Score 568; DB 22; Length 303;
Best Local Similarity 40.7%; Pred. No. 3.7e-48;
Matches 121; Conservative 48; Mismatches 116; Indels 12; Gaps 5;

```
Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDK--LVSFIIPLAGSPGRFVV 73
              :|||||||  : |:||:||       ::|   | |    |::     |||:|:   |  |' |
Db         15 LGEGPHWDVDRQSLYYVDLESAGINRYDFKQNKVYRAKIEGEIFASFILPVENKPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
                    |: ||||||       :  |:| :| ||||| | || |  : |||| |:|
Db         75 GCGLRTVIVQWDGVSAVAKVTRTLFEVQPDLKENRLNDAKTDPNGRFYGGTMA-DSGDIF 133

Qy        134 GPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSI 193
                  | ||   |  :     | : |:|||||    || |:||:    |  |||: || ::
Db        134 TQWKGELYSWQAGGQPNAIRSKVGISNGLAWDVKAKKFYFIDTNNHEVLAYDYNQSTGAV 193

Qy        194 SNQRPLFTFEK--HEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPD 251
              || : :|   |   | | :||| |:| :||::|| | |  : |::    ::||: :|||
Db        194 SNPKVIFDLRKIRPEGPLFPDGMTVDTDGNIYVATFNGGTVFKVNPSTGKILLE-IKIPT 252

Qy        252 PQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVKL 308
               |:|||||||||| |:||:|      :  |:    |  ::||||  ||:||| :|:
Db        253 TQITSVAFGGPNLDILYVTTA------NKFDQPKPAGTTFQVTGLNAKGYAGVNLKI 303
```

RESULT 5
ABB58550
ID ABB58550 standard; Protein; 303 AA.
XX
AC ABB58550;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 2442.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.

```
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02653.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 2442; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   303 AA;

  Query Match             34.1%;  Score 550;  DB 22;  Length 303;
  Best Local Similarity   39.2%;  Pred. No. 2.3e-46;
  Matches  118;  Conservative   48;  Mismatches  115;  Indels   20;  Gaps    6;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSFIIPLAGSPGRFVV 73
              :|||||||   |:||:||    :  :|  :| |    |:  : |  |::|: | |  | |
Db         15 LGEGPHWDVARQSLYYVDLEAGSLLRYDYAQNKVYKTKIEGETLAGFVLPVEGRPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
                | : |: ||||| :   :  :  |:| :: ||||| | || |  : |||       :
Db         75 GCGRRVVIVNWDGVSPSAKVVRTLFEVQPLMEKNRLNDAKVDPRGRFFGGTMRY-----I 129

Qy        134 GPV----TGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDAS 189
              |       | ||   |  :| : : :: |:||||||    || ||||:    |  ||||
Db        130 GDEFEFRHGELYRWEAGGQVSVIKGDVGISNGLAWDEKAKKFYYIDTTDYEVKSYDYDFE 189

Qy        190 TLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTV 247
              |   || : :|   |:    :  ||| ||| ||||:|| | |  | |::    ::||: :
Db        190 TGVASNPKVIFNLRKNSPKDHLLPDGLTIDTEGNLYVATFNGATIYKVNPNTGKILLE-I 248

Qy        248 KIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVK 307
              | |  |:|| |||||||| |:||:|      :  |:    |  |:|||  |: || :|
Db        249 KFPTKQITSAAFGGPNLDILYVTTA------AKFDQPAPAGTTYKVTGLNATGYPGVNLK 302

Qy        308 L 308
              :
Db        303 V 303


RESULT 6
ABB58551
ID   ABB58551 standard; Protein; 303 AA.
XX
AC   ABB58551;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 2445.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02654.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 2445; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   303 AA;

  Query Match             34.1%;  Score 550;  DB 22;  Length 303;
  Best Local Similarity   39.2%;  Pred. No. 2.3e-46;
  Matches  118;  Conservative   48;  Mismatches  115;  Indels   20;  Gaps    6;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSFIIPLAGSPGRFVV 73
              :|||||||   |:||:||    :  :|  :| |    |:  : |  |::|: | |  | |
Db         15 LGEGPHWDVARQSLYYVDLEAGSLLRYDYAQNKVYKTKIEGETLAGFVLPVEGRPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
                | : |: ||||| :   :  :  |:| :: ||||| | || |  : |||       :
Db         75 GCGRRVVIVNWDGVSPSAKVVRTLFEVQPLMEKNRLNDAKVDPRGRFFGGTMRY-----I 129

Qy        134 GPV----TGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDAS 189
              |       | ||   |  :| : : :: |:||||||    || ||||:    |  ||||
Db        130 GDEFEFRHGELYRWEAGGQVSVIKGDVGISNGLAWDEKAKKFYYIDTTDYEVKSYDYDFE 189

Qy        190 TLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTV 247
              |   || : :|   |:    :  ||| ||| ||||:|| | |  | |::    ::||: :
Db        190 TGVASNPKVIFNLRKNSPKDHLLPDGLTIDTEGNLYVATFNGATIYKVNPNTGKILLE-I 248

Qy        248 KIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVK 307
              | |  |:|| |||||||| |:||:|      :  |:    |  |:|||  |: || :|
Db        249 KFPTKQITSAAFGGPNLDILYVTTA------AKFDQPAPAGTTYKVTGLNATGYPGVNLK 302

Qy        308 L 308
              :
Db        303 V 303


RESULT 7
ABB66692
ID   ABB66692 standard; Protein; 303 AA.
XX
AC   ABB66692;
XX
DT   26-MAR-2002  (first entry)
XX
```

DE Drosophila melanogaster polypeptide SEQ ID NO 26868.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL10795.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 26868; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 303 AA;

Query Match 34.1%; Score 550; DB 22; Length 303;
Best Local Similarity 39.2%; Pred. No. 2.3e-46;
Matches 118; Conservative 48; Mismatches 115; Indels 20; Gaps 6;

```
Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSFIIPLAGSPGRFVV 73
              :|||||||   |:||:||    :  :|  :| |    |:  : |  |::|: | |  | |
Db         15 LGEGPHWDVARQSLYYVDLEAGSLLRYDYAQNKVYKTKIEGETLAGFVLPVEGRPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
                 | : |: ||||| :   :  :  |:| :: ||||| | ||  |  : |||       :
Db         75 GCGRRVVIVNWDGVSPSAKVVRTLFEVQPLMEKNRLNDAKVDPRGRFFGGTMRY-----I 129

Qy        134 GPV----TGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDAS 189
              |       | ||   |  :| : : :: |:|||||    || ||||:    |  ||||
Db        130 GDEFEFRHGELYRWEAGGQVSVIKGDVGISNGLAWDEKAKKFYYIDTTDYEVKSYDYDFE 189

Qy        190 TLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTV 247
              |   || : :|   |:    :  ||| ||| ||||:|| | |  | |::    ::||: :
Db        190 TGVASNPKVIFNLRKNSPKDHLLPDGLTIDTEGNLYVATFNGATIYKVNPNTGKILLE-I 248

Qy        248 KIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVK 307
              | |  |:|| |||||||| |:||:|      :  |:    |  |:||||   |: || :|
Db        249 KFPTKQITSAAFGGPNLDILYVTTA------AKFDQPAPAGTTYKVTGLNATGYPGVNLK 302

Qy        308 L 308
              :
Db        303 V 303
```

RESULT 8

ABB83787
ID ABB83787 standard; Protein; 299 AA.
XX
AC ABB83787;
XX
DT 14-AUG-2002 (first entry)
XX
DE Mouse SMP30 SEQ ID NO 1.
XX
KW Mouse; senescence marker protein 30; cancer; aging; senescence; SMP30;
KW cytostatic.
XX
OS Mus musculus.
XX
PN WO200234041-A1.
XX
PD 02-MAY-2002.
XX
PF 22-OCT-2001; 2001WO-JP09243.
XX
PR 23-OCT-2000; 2000JP-0322234.
XX
PA (MARU/) MARUYAMA N.
PA (KASA/) KASAHARA Y.
XX
PI Maruyama N, Kasahara Y;
XX
DR WPI; 2002-405445/43.
DR N-PSDB; ABL99968.
XX
PT Non-human animals with defective senescence marker protein 30 function,
PT useful for producing antibody, histopathological diagnosis and onset
PT mechanism of disease due to aging, and in screening side-effects during
PT drug development -
XX
PS Disclosure; Page 36-37; 44pp; Japanese.
XX
CC The invention relates to a senescence marker protein 30 (SMP30) non-human
CC animal having a defect in the SMP30 function. The animals are useful for
CC producing antibody, evaluating its biological function, histopathological
CC diagnosis and onset mechanism of disease due to aging, and in screening
CC side-effects during development of drug and cosmetics including for
CC cancer treatment.
XX
SQ Sequence 299 AA;

Query Match 28.3%; Score 457.5; DB 23; Length 299;
Best Local Similarity 37.8%; Pred. No. 4e-37;
Matches 112; Conservative 40; Mismatches 121; Indels 23; Gaps 7;

```
Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              |  || | |:  :|:| |||   |  ::    :    ||  || :    | :|
Db         14 YRCGESPVWEEASQSLLFVDIPSKIICRWDTVSNQVQRVAVDAPVSSV--ALRQLGGYVA 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::  :   | |:       |:  :  |:   |||| |||| || |  : |||| :
Db         72 TIGTKFCALNWEN-----QSVFVLAMVDEDKKNNRFNDGKVDPAGRYFAGTMAEE----T 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
                |      |||| |  |  || :   : |:||| || | |  |||||    || :|||
Db        123 APAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
              |  |||:| ::  || |    |||  || || |||| : | |:|::    ||    | |
Db        183 QTGQISNRRIVYKMEKDE--QIPDGMCIDAEGKLWVACYNGGRVIRL---DPETGKRLQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||| :  |::|| |   |:   | :    |:::::|||||| |
Db        238 VKLPVDKTTSCCFGGKDYSEMYVTCARDGLNAEGLLRQPDAGNIFKITGLGVKGIA 293
```

RESULT 9
AAR75416

```
ID   AAR75416 standard; Protein; 299 AA.
XX
AC   AAR75416;
XX
DT   31-JAN-1996  (first entry)
XX
DE   Rat regucalcin, a calcium binding protein.
XX
KW   Calcium binding protein; regucalcin; rat.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 95
FT                   /note= "corresponds to CAA codon"
FT   Misc-difference 141
FT                   /note= "corresponds to GAC codon"
XX
PN   JP07123985-A.
XX
PD   16-MAY-1995.
XX
PF   09-NOV-1993;   93JP-0279349.
XX
PR   09-NOV-1993;   93JP-0279349.
XX
PA   (DAII-) DAIICHI KAKAGU YAKUHIN KK.
PA   (YAMA/) YAMAGUCHI M.
XX
DR   WPI; 1995-211628/28.
DR   N-PSDB; AAQ87295.
XX
PT   DNA encoding regucalcin - useful for the recombinant production of
PT   regucalcin, a calcium binding protein
XX
PS   Claim 1; Page 5-6; 6pp; Japanese.
XX
CC   A rat liver cDNA clone coding for a calcium binding protein of mol.
CC   wt. 33388 was isolated and sequenced (AAQ87295). The protein was
CC   designated "regucalcin" (AAR75416).
XX
SQ   Sequence   299 AA;

  Query Match             28.1%;  Score 454.5;  DB 16;  Length 299;
  Best Local Similarity   37.5%;  Pred. No. 8e-37;
  Matches  111;  Conservative   41;  Mismatches  121;  Indels   23;  Gaps    7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              |  || | |:  :: | |||   ||  ::    :      ||  || :       | :|
Db         14 YRCGESPVWEEASKCLLFVDIPSKTVCRWDSISNRVQRVGVDAPVSSV--ALRQSGGYVA 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::  :   | |:       |:  :  |:   |||| |||| || |  : |||| :
Db         72 TIGTKFCALNWE-----DQSVFILAMVDEDKKNNRFNDGKVDPAGRYFAGTMAEE----T 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
               |       |||| |  |  || : : : |:||| || | |  |||||    || :|||
Db        123 APAVLERHQGSLYSLFPDHSVKKYFNQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
               |  |||:| ::  || |    |||  || || |||| : | |:|::    ||    | |
Db        183 PTGQISNRRTVYKMEKDE--QIPDGMCIDVEGKLWVACYNGGRVIRL---DPETGKRLQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||| :  |::|| |   :    | :    |:::::|||||| |
Db        238 VKLPVDKTTSCCFGGKDYSEMYVTCARDGMSAEGLLRQPDAGNIFKITGLGVKGIA 293


RESULT 10
AAR74219
ID   AAR74219 standard; Protein; 299 AA.
XX
AC   AAR74219;
XX
DT   03-JAN-1996  (first entry)
XX
DE   Human ageing marker protein, SMP30.
XX
KW   SMP30; marker protein; ageing; organ development.
XX
OS   Homo sapiens.
XX
PN   JP07097399-A.
XX
PD   11-APR-1995.
XX
PF   29-SEP-1993;  93JP-0265681.
XX
PR   29-SEP-1993;  93JP-0265681.
XX
PA   (FJRE ) FUJI REBIO KK.
XX
DR   WPI; 1995-175363/23.
DR   N-PSDB; AAQ90035.
XX
PT   Novel polypeptide for detecting human ageing marker protein SMP30 -
PT   for monitoring liver and kidney development in new-born babies
XX
PS   Claim 1; Page 7-8; 10pp; Japanese.
XX
CC   AAQ90035 is the SMP30 gene. It encodes the human ageing marker protein,
CC   SMP30 (AAR74219). Human SMP30 is found in human organs, tissues, blood,
CC   urine and cerebrospinal fluid. The blood concentration of SMP30 is
CC   known to increase with renal and hepatic deficiencies and to decrease
CC   with age. It is therefore useful in the monitoring of renal or hepatic
CC   deficiencies and for the monitoring of the development of the liver
CC   and kidneys in newborn babies.
XX
SQ   Sequence   299 AA;

  Query Match             27.8%;  Score 448.5;  DB 16;  Length 299;
  Best Local Similarity   36.9%;  Pred. No. 3.2e-36;
  Matches  110;  Conservative   42;  Mismatches  113;  Indels   33;  Gaps    8;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKY---TFCKVDKLVSFIIPLAG----SPG 69
              || | |:  : :| |||         :|::|     :| |  : |:   |::       |
Db         17 GESPVWEEVSNSLLFVD---------IPAKKVCRWDSFTKQVQRVTMDAPVSSVALRQSG 67

Qy         70 RFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDA 129
               :| ::  :   | |   ||       :  |:   |||| |||| || |  : |||| :
Db         68 GYVATIGTKFCALNWKEQSAV-----VLATVDNDKKNNRFNDGKVDPAGRYFAGTMAEE- 121

Qy        130 GLPVGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEY 184
                    |       |:|| |  |  || :   : |:||| || | |  |||||    || :
Db        122 ---TAPAVLERHQGALYSLFPDHHVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYSVDAF 178

Qy        185 DYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLL 244
              |||  |  |||:| ::  || |    |||  || || |||| : | |:|::     :  |
Db        179 DYDLQTGQISNRRSVYKLEKEE--QIPDGMCIDAEGKLWVACYNGGRVIRLDPVTGK-RL 235

Qy        245 DTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              |||:|  : ||  ||| |  |::|| |   :|   | :    | ::::|||||| |
Db        236 QTVKLPVDKTTSCCFGGKNYSEMYVTCARDGMDPEGLLRQPEAGGIFKITGLGVKGIA 293


RESULT 11
AAW14475
ID   AAW14475 standard; peptide; 299 AA.
XX
AC   AAW14475;
XX
DT   28-MAY-1997  (first entry)
XX
DE   Human senility marker protein, hSMP30.
XX
KW   SMP30; senility marker protein; monoclonal antibody; detection.
```

```
XX
OS   Homo sapiens.
XX
PN   JP08319298-A.
XX
PD   03-DEC-1996.
XX
PF   25-MAY-1995;   95JP-0149791.
XX
PR   25-MAY-1995;   95JP-0149791.
XX
PA   (FJRE ) FUJI REBIO KK.
XX
DR   WPI; 1997-073109/07.
XX
PT   Anti-human senility marker protein monoclonal antibody - useful for
PT   detection of protein
XX
PS   Claim 3; Page 6-7; 8pp; Japanese.
XX
CC   The present sequence is that of human senility marker protein (hSMP30)
CC   which has a molecular weight of 30 kDa. Monoclonal antibodies recognising
CC   hSMP30 are claimed and can be used in a method for detection of the
CC   hSMP30 protein in a sample.
XX
SQ   Sequence   299 AA;

  Query Match             27.8%;  Score 448.5;  DB 18;  Length 299;
  Best Local Similarity   36.9%;  Pred. No. 3.2e-36;
  Matches  110;  Conservative   42;  Mismatches  113;  Indels   33;  Gaps    8;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKY---TFCKVDKLVSFIIPLAG----SPG 69
              || | |:  : :| |||          :|::|     :| |  : |:   |::       |
Db         17 GESPVWEEVSNSLLFVD---------IPAKKVCRWDSFTKQVQRVTMDAPVSSVALRQSG 67

Qy         70 RFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDA 129
               :| ::  :   | |   ||       :  |:   |||| |||| || |  : |||| :
Db         68 GYVATIGTKFCALNWKEQSAV-----VLATVDNDKKNNRFNDGKVDPAGRYFAGTMAEE- 121

Qy        130 GLPVGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEY 184
                 |       |:|| |  | || :   : |:||| || | |  |||||    || :
Db        122 ---TAPAVLERHQGALYSLFPDHHVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYSVDAF 178

Qy        185 DYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLL 244
              |||  |  |||:| ::  || |    |||  || || |||| : | |:|::     :  |
Db        179 DYDLQTGQISNRRSVYKLEKEE--QIPDGMCIDAEGKLWVACYNGGRVIRLDPVTGK-RL 235

Qy        245 DTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              |||:|  : ||  ||| |  |::|| |   :|   | :    | ::::|||||| |
Db        236 QTVKLPVDKTTSCCFGGKNYSEMYVTCARDGMDPEGLLRQPEAGGIFKITGLGVKGIA 293


RESULT 12
AAY06995
ID   AAY06995 standard; Protein; 299 AA.
XX
AC   AAY06995;
XX
DT   02-JUL-1999  (first entry)
XX
DE   Senescence marker protein SMP30.
XX
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer.
XX
OS   Homo sapiens.
XX
PN   WO9904265-A2.
XX
PD   28-JAN-1999.
XX
PF   15-JUL-1998;   98WO-US14679.
XX
PR   22-JUN-1998;   98US-0102322.
PR   17-JUL-1997;   97US-0896164.
PR   10-OCT-1997;   97US-0061599.
PR   10-OCT-1997;   97US-0061765.
PR   10-OCT-1997;   97US-0948705.
PR   11-OCT-1997;   97GB-0021697.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Chen Y,  Gout I,  Gure A,  O'Hare M,  Obata Y,  Old LJ;
PI   Pfreundschuh M,  Sahin U,  Scanlan MJ,  Stockert E;
PI   Tureci O;
XX
DR   WPI; 1999-132448/11.
DR   N-PSDB; AAX40196.
XX
PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
XX
PS   Example 8; Page 775-776; 787pp; English.
XX
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
XX
SQ   Sequence   299 AA;

  Query Match             27.8%;  Score 448.5;  DB 20;  Length 299;
  Best Local Similarity   36.9%;  Pred. No. 3.2e-36;
  Matches  110;  Conservative   42;  Mismatches  113;  Indels   33;  Gaps    8;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKY---TFCKVDKLVSFIIPLAG----SPG 69
              || | |:  : :| |||          :|::|     :| |  : |:   |::       |
Db         17 GESPVWEEVSNSLLFVD---------IPAKKVCRWDSFTKQVQRVTMDAPVSSVALRQSG 67

Qy         70 RFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDA 129
               :| ::  :   | |   ||       :  |:   |||| |||| || |  : |||| :
Db         68 GYVATIGTKFCALNWKEQSAV-----VLATVDNDKKNNRFNDGKVDPAGRYFAGTMAEE- 121

Qy        130 GLPVGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEY 184
                 |       |:|| |  | || :   : |:||| || | |  |||||    || :
Db        122 ---TAPAVLERHQGALYSLFPDHHVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYSVDAF 178

Qy        185 DYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLL 244
              |||  |  |||:| ::  || |    |||  || || |||| : | |:|::     :  |
Db        179 DYDLQTGQISNRRSVYKLEKEE--QIPDGMCIDAEGKLWVACYNGGRVIRLDPVTGK-RL 235

Qy        245 DTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              |||:|  : ||  ||| |  |::|| |   :|   | :    | ::::|||||| |
Db        236 QTVKLPVDKTTSCCFGGKNYSEMYVTCARDGMDPEGLLRQPEAGGIFKITGLGVKGIA 293


RESULT 13
AAB11651
ID   AAB11651 standard; Protein; 293 AA.
XX
AC   AAB11651;
XX
DT   23-OCT-2000  (first entry)
XX
```

```
DE   A. vitis hypersensitive response elicitor protein, SEQ ID NO:29.
XX
KW   Hypersensitive response elicitor protein; HR; disease resistance;
KW   insecticide; fungicide; antiviral; bactericide; growth enhancer;
KW   stress resistance; transgenic plant.
XX
OS   Agrobacterium vitis.
XX
PN   WO200028056-A2.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26079.
XX
PR   06-NOV-1998;   98US-0107387.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Burr TJ,  Herlache TC,  Zhang H;
XX
DR   WPI; 2000-376567/32.
DR   N-PSDB; AAA61508.
XX
PT   New protein from Agrobacterium vitis, useful e.g. for imparting
PT   resistance to disease or stress to plants, is involved in production of
PT   a hypersensitive response -
XX
PS   Claim 3; Page 104-105; 157pp; English.
XX
CC   Sequences AAB11630-B11688 represent proteins from Agrobacterium vitis
CC   which elicit a hypersensitivity response (HR) in a plant. The invention
CC   also relates to nucleotide sequences (AAA61501-A61524) encoding the A.
CC   vitis HR elicitor proteins. The HR is a rapid, localised necrosis that
CC   is associated with the active defence of plants against many pathogens,
CC   and occurs when a pathogenic organism interacts with a nonhost plant
CC   (i.e. one in which intracellular bacterial growth and disease development
CC   do not occur). Like other HR elicitors, the A. vitis elicitor functions
CC   in non-host plants by causing a rapid hypersensitive response that
CC   results in walling-off and killing of the pathogen. On grape plants, the
CC   A. vitis elicitor induces a restricted necrosis of tissues, resulting in
CC   the death of plant cells and induction of pathogen resistance. A. vitis
CC   HR elicitor proteins, in non-infectious form, are used to treat plants or
CC   their seeds to impart resistance to disease, such as those caused by
CC   fungi, bacteria or viruses; and to enhance growth, e.g., to increase
CC   yield or to provide earlier germination or maturation. The proteins can
CC   also be used to control insects, to impart resistance to environmental
CC   stresses, e.g., cold, and to improve nutritional value, e.g., altered oil
CC   content. The same effects can be produced by producing transgenic plants
CC   or seeds by incorporation of DNA that encodes A. vitis HR elicitor
CC   proteins. Use of A. vitis HR elicitor proteins, or nucleic acids encoding
CC   them, may allow control of previously untreatable diseases; provide
CC   systemic treatment; and eliminate the need for biological control agents
CC   or polluting chemicals.
XX
SQ   Sequence   293 AA;

  Query Match             14.7%;  Score 238;  DB 21;  Length 293;
  Best Local Similarity   26.5%;  Pred. No. 3.4e-15;
  Matches   80;  Conservative   46;  Mismatches  126;  Indels   50;  Gaps   12;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTF--HKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              :|||| :|    | :: | : |    |::   :           :   :|:  |     :
Db         17 LGEGPTYDVARDTAWWFDILGKGLIEHRFATGE----------TIRHDLPMMAS-ALATI 65

Qy         74 SLEREIAILTWDGV---SAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAG 130
                ||:: : |  |:   : |   :  :  :||    || |||:  | | || |||    |
Db         66 DAERQM-LATEQGIFIRTIASGELTLLTALEPERPGNRSNDGRVHPCGALWIGTM----G 120

Qy        131 LPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDAST 190
                  |::||:  ||  ::::| |:| | : :| |    ||:|:    : :   | :|
Db        121 KTAADGAGAIYHVAGDKVTRLYDS-ISIPNAICFSPDGSLGYYVDTRVNTLMKVSLDPAT 179

Qy        191 LSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVF----------QGQRIIKISTQQP 240
                :    |    :     |  ||   | || || | :           || : :
Db        180 -GLPTGMPTVHIDGRGKDGGIDGAVCDAEGGLWNARWGVGAVDHYDRTGQHLAR------ 232

Qy        241 EVLLDTVKIPDPQVTSVAFGGPNLDELHVTSA--GLQLDDSSLDKSLVNGHVYRVTGLGV 298
                    ::|  | |  || |   | | ||||  ||  |  : |    :|    |  : |
Db        233 ------YRLPAAQTTCPAFIGAKADRLLVTSATEGLDADGLAADP---HGGKTLVLDITV 283

Qy        299 KG 300
               ||
Db        284 KG 285


RESULT 14
AAB11650
ID   AAB11650 standard; Protein; 340 AA.
XX
AC   AAB11650;
XX
DT   23-OCT-2000  (first entry)
XX
DE   A. vitis hypersensitive response elicitor protein, SEQ ID NO:28.
XX
KW   Hypersensitive response elicitor protein; HR; disease resistance;
KW   insecticide; fungicide; antiviral; bactericide; growth enhancer;
KW   stress resistance; transgenic plant.
XX
OS   Agrobacterium vitis.
XX
PN   WO200028056-A2.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;   99WO-US26079.
XX
PR   06-NOV-1998;   98US-0107387.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Burr TJ,  Herlache TC,  Zhang H;
XX
DR   WPI; 2000-376567/32.
DR   N-PSDB; AAA61508.
XX
PT   New protein from Agrobacterium vitis, useful e.g. for imparting
PT   resistance to disease or stress to plants, is involved in production of
PT   a hypersensitive response -
XX
PS   Claim 3; Page 102-104; 157pp; English.
XX
CC   Sequences AAB11630-B11688 represent proteins from Agrobacterium vitis
CC   which elicit a hypersensitivity response (HR) in a plant. The invention
CC   also relates to nucleotide sequences (AAA61501-A61524) encoding the A.
CC   vitis HR elicitor proteins. The HR is a rapid, localised necrosis that
CC   is associated with the active defence of plants against many pathogens,
CC   and occurs when a pathogenic organism interacts with a nonhost plant
CC   (i.e. one in which intracellular bacterial growth and disease development
CC   do not occur). Like other HR elicitors, the A. vitis elicitor functions
CC   in non-host plants by causing a rapid hypersensitive response that
CC   results in walling-off and killing of the pathogen. On grape plants, the
CC   A. vitis elicitor induces a restricted necrosis of tissues, resulting in
CC   the death of plant cells and induction of pathogen resistance. A. vitis
CC   HR elicitor proteins, in non-infectious form, are used to treat plants or
CC   their seeds to impart resistance to disease, such as those caused by
CC   fungi, bacteria or viruses; and to enhance growth, e.g., to increase
CC   yield or to provide earlier germination or maturation. The proteins can
CC   also be used to control insects, to impart resistance to environmental
CC   stresses, e.g., cold, and to improve nutritional value, e.g., altered oil
CC   content. The same effects can be produced by producing transgenic plants
CC   or seeds by incorporation of DNA that encodes A. vitis HR elicitor
CC   proteins. Use of A. vitis HR elicitor proteins, or nucleic acids encoding
CC   them, may allow control of previously untreatable diseases; provide
CC   systemic treatment; and eliminate the need for biological control agents
CC   or polluting chemicals.
```

```
XX
SQ   Sequence   340 AA;

  Query Match            14.7%;  Score 238;  DB 21;  Length 340;
  Best Local Similarity   26.5%;  Pred. No. 4.2e-15;
  Matches   80;  Conservative   46;  Mismatches  126;  Indels   50;  Gaps   12;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTF--HKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              :|||| :|    | :: | : |    |::   :           :   :|:  |     :
Db         64 LGEGPTYDVARDTAWWFDILGKGLIEHRFATGE----------TIRHDLPMMAS-ALATI 112

Qy         74 SLEREIAILTWDGV---SAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAG 130
                ||:: : |  |:   : |   :  :  :||    || |||:  | | || |||    |
Db        113 DAERQM-LATEQGIFIRTIASGELTLLTALEPERPGNRSNDGRVHPCGALWIGTM----G 167

Qy        131 LPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDAST 190
                    |::||:  ||  ::::| |:| | : :| |    ||:|:    : :   | :|
Db        168 KTAADGAGAIYHVAGDKVTRLYDS-ISIPNAICFSPDGSLGYYVDTRVNTLMKVSLDPAT 226

Qy        191 LSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVF----------QGQRIIKISTQQP 240
              :   |    :    |  ||   | || || | :           || : :
Db        227 -GLPTGMPTVHIDGRGKDGGIDGAVCDAEGGLWNARWGVGAVDHYDRTGQHLAR------ 279

Qy        241 EVLLDTVKIPDPQVTSVAFGGPNLDELHVTSA--GLQLDDSSLDKSLVNGHVYRVTGLGV 298
                  ::|  | |  || |   | | ||||  ||  |  : |    :|    |  : |
Db        280 ------YRLPAAQTTCPAFIGAKADRLLVTSATEGLDADGLAADP---HGGKTLVLDITV 330

Qy        299 KG 300
              ||
Db        331 KG 332


RESULT 15
ABP32045
ID   ABP32045 standard; Protein; 107 AA.
XX
AC   ABP32045;
XX
DT   09-JUL-2002  (first entry)
XX
DE   Human RNA polymerase-like ORF1018 protein, SEQ ID NO:2036.
XX
KW   Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW   disease monitoring; cytokine; cell proliferation; cell differentiation;
KW   immune modulation; haematopoiesis regulation; tissue growth;
KW   angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW   behaviour; cancer; proliferative disorder; neurological disorder;
KW   cardiovascular disease; immune system disorder; organ transplantation;
KW   tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW   hypothyroidism; cholesterol ester storage disease; infection; vulnerary;
KW   vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW   neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW   cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW   dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS   Homo sapiens.
XX
PN   WO200190366-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US17076.
XX
PR   24-MAY-2000; 2000US-206690P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Leach MD,  Shimkets RA;
XX
DR   WPI; 2002-106200/14.
DR   N-PSDB; ABN76071.
XX
PT   Novel human polypeptides and polynucleotides useful for diagnosing,
PT   preventing and treating cardiovascular disease, neurodegenerative,
PT   hyperproliferative disorders and disorders related to organ
PT   transplantation -
XX
PS   Claim 10; Page 773; 2508pp; English.
XX
CC   Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC   designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC   ABN79587 represent cDNAs encoding them. The invention also encompasses
CC   polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC   referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC   the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC   polynucleotides, the recombinant production of ORFX proteins, antibodies
CC   specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC   polypeptides, methods of screening for modulators of ORFX expression or
CC   activity, and methods of screening individuals for a predisposition to an
CC   ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC   range of biological activities, such as cytokine, cell proliferation,
CC   cell differentiation, immune modulation, haematopoiesis regulation,
CC   tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC   chemokinetic activity, haemostatic activity, thrombolytic activity,
CC   receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC   and antiinfective activity, and may also be involved in the determination
CC   of bodily characteristics, fertility and behaviour. ORFX proteins,
CC   nucleic acids and antibodies may be used in the treatment of cancers,
CC   other proliferative disorders such as psoriasis and benign tumours,
CC   neurological disorders such as epilepsy and Alzheimer's disease,
CC   cardiovascular diseases, immune system disorders, disorders related to
CC   organ transplantation, disorders of tissue growth and regeneration,
CC   diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC   storage disease, and infectious diseases caused by viral, bacterial,
CC   fungal and other pathogens. ORFX nucleic acids may also be used as a
CC   source of primers and probes, in the detection of ORFX genomic sequences
CC   or transcripts, in the identification and cloning of homologous
CC   sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC   nucleic acids may additionally be used to produce transgenic animals
CC   which may be useful for studying the function and/or activity of ORFX
CC   protein, and in drug screening. The ORFX proteins may also be used as
CC   immunogens to generate specific antibodies, which are useful in the
CC   diagnosis, treatment and monitoring of ORFX-associated diseases.
XX
SQ   Sequence   107 AA;

  Query Match            12.1%;  Score 195.5;  DB 23;  Length 107;
  Best Local Similarity   43.4%;  Pred. No. 1.3e-11;
  Matches   49;  Conservative    9;  Mismatches   44;  Indels   11;  Gaps    3;

Qy        115 DPLGNLWTGTMAIDAGLPVGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLK 169
              || |  : |||| :     |       |||| |  |  || :   : |:||| || | |
Db          1 DPAGRYFAGTMAEE----TAPAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSLDHK 56

Qy        170 KMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNL 222
                |||||    |  :|||  |  |||:| ::  || |    |||  || || |
Db         57 IFYYIDSLSYTVXAFDYDLPTGQISNRRTVYKMEKDE--QIPDGMCIDVEGKL 107


Search completed: August  1, 2003, 14:32:07
Job time : 67 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:31:08 ; Search time 37 Seconds
(without alignments)
800.254 Million cell updates/sec

Title: US-10-089-986-2
Perfect score: 1615
Sequence: 1 MGPVVEKIAELGKYTVGEGP.....HVYRVTGLGVKGFAGVKVKL 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	550	34.1	303	2	JC7250	anterior fat body
2	538	33.3	306	2	JC7249	anterior fat body
3	457.5	28.3	299	2	S72173	senescence marker
4	454.5	28.1	299	2	S34588	senescence marker
5	448.5	27.8	299	1	S60035	senescence marker
6	423.5	26.2	569	2	E95309	probable calcium b
7	409.5	25.4	560	2	AI3216	hypothetical prote
8	376	23.3	292	1	A70047	RNA polymerase hom
9	365.5	22.6	285	2	B90445	regucalcin homolog
10	331	20.5	289	2	A87351	Smp-30/Cgr1 family
11	315	19.5	293	2	A87473	Smp-30/Cgr1 family
12	292	18.1	285	2	G90485	conserved hypothet
13	288.5	17.9	311	2	D95265	probable regucalci
14	273	16.9	293	2	AB3071	calcium-binding pr
15	273	16.9	322	2	G98215	senescence marker
16	256	15.9	295	2	AH2662	calcium-binding pr
17	256	15.9	314	2	G97444	probable calcium-b
18	206	12.8	294	2	AF3043	calcium-binding pr
19	206	12.8	305	2	F98242	regucalcin (AB0379
20	203	12.6	302	2	AH0832	conserved hypothet
21	202.5	12.5	291	2	AG3501	gluconolactonase (
22	153.5	9.5	306	2	AH3051	gluconolactonase p
23	153.5	9.5	348	2	D98234	probable gluconola
24	140	8.7	303	2	C95896	probable gluconola
25	132	8.2	304	2	G95274	probable gluconola
26	124.5	7.7	320	2	S28218	gluconolactonase (
27	120	7.4	347	2	B87401	gluconolactonase [
28	113	7.0	303	2	G87512	gluconolactonase,
29	111	6.9	356	2	D82643	conserved hypothet
30	104	6.4	324	2	JC7119	Drp35 protein - St
31	103.5	6.4	1116	2	B88612	protein Y76A2A.2 [
32	103.5	6.4	1238	1	JC5573	copper-transportin
33	102.5	6.3	480	2	G70302	conserved hypothet
34	102	6.3	324	2	H90077	Drp35 [imported] -
35	101	6.3	1376	2	F83711	exopolygalacturona
36	100	6.2	358	2	S45911	hypothetical prote
37	100	6.2	371	2	T45775	mucin-like protein
38	99	6.1	338	2	B82698	gluconolactonase p
39	98	6.1	384	2	A72257	probable transamin
40	97.5	6.0	553	2	A81057	hypothetical integ
41	97.5	6.0	743	2	AH2060	hypothetical prote
42	97	6.0	371	2	T45776	mucin-like protein
43	96	5.9	5198	2	T43290	hemicentin precurs
44	95.5	5.9	954	2	AI0438	probable exported
45	94.5	5.9	409	2	T07239	translation elonga

ALIGNMENTS

RESULT 1
JC7250
anterior fat body protein - fruit fly (Drosophila sp.)
C;Species: Drosophila sp.
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 17-Nov-2000
C;Accession: JC7250
R;Nakajima, Y.; Natori, S.
J. Biochem. 127, 901-908, 2000
A;Title: Identification and characterization of an anterior fat body protein in an in
A;Reference number: JC7249
A;Accession: JC7250
A;Molecule type: mRNA
A;Residues: 1-303 <NAK>
A;Cross-references: DDBJ:AB036904
C;Comment: This protein, a homolog of the senescence marker protein-30, has its funct
C;Genetics:
A;Gene: afp
A;Cross-references: FlyBase:FBgn0030362
C;Superfamily: senescence marker protein-30

Query Match 34.1%; Score 550; DB 2; Length 303;
Best Local Similarity 39.2%; Pred. No. 6.5e-37;
Matches 118; Conservative 48; Mismatches 115; Indels 20; Gaps 6;

```
Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSFIIPLAGSPGRFVV 73
              :|||||||   |:||:||    :  :|  :| |    |:  : |  |::|: | |  | |
Db         15 LGEGPHWDVARQSLYYVDLEAGSLLRYDYAQNKVYKTKIEGETLAGFVLPVEGRPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
                | : |: ||||| :   :  :  |:| :: ||||| | || |  : |||       :
Db         75 GCGRRVVIVNWDGVSPSAKVVRTLFEVQPLMEKNRLNDAKVDPRGRFFGGTMRY-----I 129

Qy        134 GPV----TGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDAS 189
             |      | ||   |  :| : : :: |:|||||    || ||||:    |  ||||
Db        130 GDEFEFRHGELYRWEAGGQVSVIKGDVGISNGLAWDEKAKKFYYIDTTDYEVKSYDYDFE 189

Qy        190 TLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTV 247
             |   || : :|   |:    :  ||| ||| ||||:|| | |  | |::    ::||: :
Db        190 TGVASNPKVIFNLRKNSPKDHLLPDGLTIDTEGNLYVATFNGATIYKVNPNTGKILLE-I 248

Qy        248 KIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVK 307
             | |  |:|| |||||||| |:||:|     :  |:    |  |:||||   |: || :|
Db        249 KFPTKQITSAAFGGPNLDILYVTTA------AKFDQPAPAGTTYKVTGLNATGYPGVNLK 302

Qy        308 L 308
             :
Db        303 V 303
```

RESULT 2
JC7249
anterior fat body protein - flesh fly (Sarcophaga peregrina)

```
C;Species: Sarcophaga peregrina
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C;Accession: JC7249; PC7071
R;Nakajima, Y.; Natori, S.
J. Biochem. 127, 901-908, 2000
A;Title: Identification and characterization of an anterior fat body protein in an insec
A;Reference number: JC7249
A;Accession: JC7249
A;Molecule type: mRNA
A;Residues: 1-306 <NAK>
A;Cross-references: DDBJ:AB036903
A;Experimental source: larval fat body
A;Accession: PC7071
A;Molecule type: protein
A;Residues: 157-168;174-185;186-195 <NA2>
C;Comment: This protein, a homolog of the senescence marker protein-30, has its function
C;Genetics:
A;Gene: afp
C;Superfamily: senescence marker protein-30

  Query Match             33.3%;  Score 538;  DB 2;  Length 306;
  Best Local Similarity   39.0%;  Pred. No. 6.1e-36;
  Matches  122;  Conservative   48;  Mismatches  123;  Indels   20;  Gaps    7;

Qy          3 PVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSF 60
              |: :  |||     |||||||  || ||:||       :|  :: |    |:  : | |
Db          7 PLPDSHAEL-----GEGPHWDVATQNLYYVDINAGKLLRYNYNENKVYKAKIEGEDLAGF 61

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNV---EPHIKNNRLNDGKADPL 117
               ||: |:  :| |   | : |: |||:|     |: :  |   :     || |||| ||
Db         62 AIPVEGTTDQFAVGSGRRVVIVQWDGISETAKVIKTLFEVQTGDERFNGNRFNDGKCDPR 121

Qy        118 GNLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSG 177
              | |: |||    |     | ||      |::: :||: |:|||||:   || ||||:
Db        122 GRLFAGTMRY-VGDEFEHRYGELYKYEKGGKIEVVKSNVGISNGLAWNEKTKKFYYIDTT 180

Qy        178 KRRVDEYDYDASTLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKI 235
                | |||||  |   :| : :|   |:    :  ||| ||| :||::|| | |  | |:
Db        181 DYEVKEYDYDFETGKATNPKVVFNLRKNSPKDHLLPDGMTIDTDGNIYVATFNGYTIYKV 240

Qy        236 STQQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTG 295
              :    :|||: :| |  |:|| |||||||| |:||::      |   :    |  |:|||
Db        241 NPTTGKVLLE-IKFPCKQITSAAFGGPNLDILYVTTS------SRFGEPDPAGTTYKVTG 293

Qy        296 LGVKGFAGVKVKL 308
              || ||:   |::|
Db        294 LGAKGYPMTKIQL 306


RESULT 3
S72173
senescence marker protein 30 - mouse
N;Alternate names: regucalcin
C;Species: Mus musculus (house mouse)
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Sep-1999
C;Accession: S72173; S72174
R;Fujita, T.; Shirasawa, T.; Maruyama, N.
Biochim. Biophys. Acta 1308, 49-57, 1996
A;Title: Isolation and characterization of genomic and cDNA clones encoding mouse senesc
A;Reference number: S72173; MUID:96328264; PMID:8765750
A;Accession: S72173
A;Molecule type: mRNA
A;Residues: 1-299 <FUJ>
A;Cross-references: EMBL:U28937; NID:g1143999; PIDN:AAC52721.1; PID:g1144000
A;Accession: S72174
A;Molecule type: DNA
A;Residues: 1-54 <FUW>
A;Cross-references: EMBL:U32170; NID:g1144337; PIDN:AAD03478.1; PID:g1144338
A;Experimental source: liver
C;Function:
A;Description: calcium binding
A;Note: age-associated decrease
C;Superfamily: senescence marker protein-30
C;Keywords: calcium binding

  Query Match             28.3%;  Score 457.5;  DB 2;  Length 299;
  Best Local Similarity   37.8%;  Pred. No. 1.9e-29;
  Matches  112;  Conservative   40;  Mismatches  121;  Indels   23;  Gaps    7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              |  || | |:  :|:| |||   |   ::      :     ||  || :       | :|
Db         14 YRCGESPVWEEASQSLLFVDIPSKIICRWDTVSNQVQRVAVDAPVSSV--ALRQLGGYVA 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::  :   | |:       |:  :  |:   |||| |||| || |  : |||| :
Db         72 TIGTKFCALNWEN-----QSVFVLAMVDEDKKNNRFNDGKVDPAGRYFAGTMAEE----T 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
               |       |||| |  |  || :   : |:||| || | |  |||||    || :|||
Db        123 APAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
               |  |||:| ::  || |    |||  || || |||| : | |:|::    ||    | |
Db        183 QTGQISNRRIVYKMEKDE--QIPDGMCIDAEGKLWVACYNGGRVIRL---DPETGKRLQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||| :  |::|| |   |:   | :    |:::::|||||| |
Db        238 VKLPVDKTTSCCFGGKDYSEMYVTCARDGLNAEGLLRQPDAGNIFKITGLGVKGIA 293


RESULT 4
S34588
senescence marker protein SMP-30 - rat
N;Alternate names: regucalcin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S34588; S34589; S27203
R;Shimokawa, N.; Yamaguchi, M.
FEBS Lett. 327, 251-255, 1993
A;Title: Molecular cloning and sequencing of the cDNA coding for a calcium-binding pr
A;Reference number: S34588; MUID:93351639; PMID:8348951
A;Accession: S34588
A;Molecule type: mRNA
A;Residues: 1-299 <SHI>
A;Cross-references: GB:D38467; GB:D14327; GB:D16386; NID:g600378; PIDN:BAA07490.1; PI
A;Accession: S34589
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 84-299 <SH2>
A;Cross-references: EMBL:D14327
R;Fujita, T.; Shirasawa, T.; Uchida, K.; Maruyama, N.
Biochim. Biophys. Acta 1132, 297-305, 1992
A;Title: Isolation of cDNA clone encoding rat senescence marker protein-30 (SMP30) an
A;Reference number: S27203; MUID:93041931; PMID:1420310
A;Accession: S27203
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-147,'D',149-299 <FUJ>
A;Cross-references: EMBL:X69021; NID:g57254; PIDN:CAA48786.1; PID:g57255
C;Superfamily: senescence marker protein-30
C;Keywords: calcium binding

  Query Match             28.1%;  Score 454.5;  DB 2;  Length 299;
  Best Local Similarity   37.5%;  Pred. No. 3.4e-29;
  Matches  111;  Conservative   41;  Mismatches  121;  Indels   23;  Gaps    7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              |  || | |:  :: | |||   ||  ::      :     ||  || :       | :|
Db         14 YRCGESPVWEEASKCLLFVDIPSKTVCRWDSISNRVQRVGVDAPVSSV--ALRQSGGYVA 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::  :   | |:       |:  :  |:   |||| |||| || |  : |||| :
Db         72 TIGTKFCALNWE-----DQSVFILAMVDEDKKNNRFNDGKVDPAGRYFAGTMAEE----T 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
               |       |||| |  |  || : : : |:||| || | |  |||||    || :|||
```

```
Db        123 APAVLERHQGSLYSLFPDHSVKKYFNQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
              |  |||:| ::  || |    |||  || || |||| : | |:|::    ||    | |
Db        183 PTGQISNRRTVYKMEKDE--QIPDGMCIDVEGKLWVACYNGGRVIRL---DPETGKRLQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||| :  |::|| |   :    | :    |:::::|||||| |
Db        238 VKLPVDKTTSCCFGGKDYSEMYVTCARDGMSAEGLLRQPDAGNIFKITGLGVKGIA 293


RESULT 5
S60035
senescence marker protein 30 - human
N;Alternate names: regucalcin
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: S60035; I52491
R;Fujita, T.; Mandel, J.L.; Shirasawa, T.; Hino, O.; Shirai, T.; Maruyama, N.
Biochim. Biophys. Acta 1263, 249-252, 1995
A;Title: Isolation of cDNA clone encoding human homologue of senescence marker protein-3
A;Reference number: I52491; MUID:96004897; PMID:7548213
A;Accession: S60035
A;Molecule type: mRNA
A;Residues: 1-299 <FUJ>
A;Cross-references: EMBL:D31815; NID:g1072311; PIDN:BAA06602.1; PID:g1072312
C;Genetics:
A;Gene: GDB:RGN; RC; SMP30
A;Cross-references: GDB:9955055
A;Map position: Xp11.3-Xp11.2
C;Superfamily: senescence marker protein-30
C;Keywords: calcium binding

  Query Match             27.8%;  Score 448.5;  DB 1;  Length 299;
  Best Local Similarity   36.9%;  Pred. No. 1e-28;
  Matches  110;  Conservative   42;  Mismatches  113;  Indels   33;  Gaps    8;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKY---TFCKVDKLVSFIIPLAG----SPG 69
              || | |:  : :| |||         :|::|     :| |  : |:   |::       |
Db         17 GESPVWEEVSNSLLFVD---------IPAKKVCRWDSFTKQVQRVTMDAPVSSVALRQSG 67

Qy         70 RFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDA 129
              :| ::  :   | |   ||       :  |:   |||| |||| || |  : |||| :
Db         68 GYVATIGTKFCALNWKEQSAV-----VLATVDNDKKNNRFNDGKVDPAGRYFAGTMAEE- 121

Qy        130 GLPVGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEY 184
                  |      |:|| |  |  || :   : |:||| || | |  |||||    || :
Db        122 ---TAPAVLERHQGALYSLFPDHHVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYSVDAF 178

Qy        185 DYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLL 244
              |||  |  |||:| ::  || |    |||  || || |||| : | |:|::     :  |
Db        179 DYDLQTGQISNRRSVYKLEKEE--QIPDGMCIDAEGKLWVACYNGGRVIRLDPVTGK-RL 235

Qy        245 DTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              |||:|  : ||  ||| |  |::|| |   :|   | :    | ::::|||||| |
Db        236 QTVKLPVDKTTSCCFGGKNYSEMYVTCARDGMDPEGLLRQPEAGGIFKITGLGVKGIA 293


RESULT 6
E95309
probable calcium binding transcription regulatory protein [imported] - Sinorhizobium mel
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: E95309
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: E95309
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-569 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65039.1; PID:g14523470; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa0717
A;Genome: plasmid

  Query Match             26.2%;  Score 423.5;  DB 2;  Length 569;
  Best Local Similarity   34.6%;  Pred. No. 2.7e-26;
  Matches  100;  Conservative   46;  Mismatches  102;  Indels   41;  Gaps    7;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75
              :|| | |    : ||:|| :    |:: |   |   | | |||| ::|          :
Db        293 LGEAPVWVEREKRLYWVDILHPAVHRFDPVTGKNESCNVAKLVSAVLP----------TR 342

Qy         76 EREIAILTWDGVS----------AAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTG 123
                 : : : |||             | |         || :  ||||| | || | || |
Db        343 NEGLIVASQDGVEHFDFDRGDFNPFAEP---------EPGLPENRLNDAKVDPSGRLWVG 393

Qy        124 TMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDE 183
              :| :|   |    ||||| | :  :|    |   :|||||| |    |::|:    :
Db        394 SMRLDVSRP----TGSLYRLTSAGEVTRAGSGFTVANGLAWSPDSSTFYFVDTVPGIIYA 449

Qy        184 YDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVL 243
              ||:||   ||:|:|   |    |  | |||  :| :| :| |:: | |   ::  :|:
Db        450 YDFDAREGSIANRRVFVTVP--EAEGRPDGLAVDADGGVWCAIWDGWR---VNRYRPDGR 504

Qy        244 LD-TVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVY 291
              ||  |::| |: |||||||  |  | :|||  :|  |:| :: ::| ::
Db        505 LDRAVELPVPRPTSVAFGGDELATLFITSARTRLPASTLTEAPLSGGIF 553


RESULT 7
AI3216
hypothetical protein Atu5464 [imported] - Agrobacterium tumefaciens (strain C58, Dupo
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AI3216
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AI3216
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-560 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL46151.1; PID:g17743920; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu5464
A;Genome: plasmid

  Query Match             25.4%;  Score 409.5;  DB 2;  Length 560;
  Best Local Similarity   33.1%;  Pred. No. 3.6e-25;
  Matches   99;  Conservative   57;  Mismatches  110;  Indels   33;  Gaps    9;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75
              :|| | |      ||:|| :    |:: |:  :   |:  |||| :||: |  || :|:
Db        284 LGESPVWHEGEDALYWVDILHPAVHRFDPATGRNETCETGKLVSAVIPVTG--GRLLVAS 341

Qy         76 EREIAILTWDGV---SAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLP 132
```

```
Db        342 Q--------DGVEWLDFASGRLTPFVSPEAGIADNRLNDAKCGPDGAIWVGSMRIDASKP 393

Qy        133 VGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLS 192

Db        394 ----TGALYRINANGASERKEGGIIVSNGLGWSPDGRTFYFVDTVPGLIHAYDCDPATGA 449

Qy        193 ISNQRPLFTFEKHEVP---GYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLD-TVK 248

Db        450 LSQRR-----EFARIPVADGRPDGLAVDAEGGVWCAIWDGWCVRR--YLPNGKLDQVID 501

Qy        249 IPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYR----VTGLGVKGFAG 303

Db        502 MPVPRPSSIAFGGPDLSTLFITSARTRLPASTLADAPLSGGLFSCRPGISGARIAMFEG 560


RESULT 8
A70047
RNA polymerase homolog yvrE - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: A70047
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A70047
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-292 <KUN>
A;Cross-references: GB:Z99120; GB:Z99121; GB:AL009126; NID:g2635827; PID:g2635833; NID:g
A;Experimental source: strain 168
C;Genetics:
A;Gene: yvrE
C;Superfamily: senescence marker protein-30

  Query Match             23.3%;  Score 376;  DB 1;  Length 292;
  Best Local Similarity   31.5%;  Pred. No. 7.2e-23;
  Matches   85;  Conservative   52;  Mismatches  105;  Indels   28;  Gaps    7;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75

Db         13 IGEGPLWDEENGRLYWVDILGSELHIFDPEE------KINRSIKF--------KSFVTAL 58

Qy         76 ER----EIAILTWDG---VSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAID 128

Db         59 AKYSKDELIMTMKDGFYLYHLRDDSLEKIKQPKDMHESLRFNDAKCDPYGRLWAGTTSME 118

Qy        129 AGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188

Db        119 G----EQKQASLYRLNLDGSLVKIKDQVSTSNGLDWDRERNLMYYIDTPTQEIVRYSYDP 174

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVK 248

Db        175 QSGDVSNPEPVYRFDQSD--GLPDGMTIDQNGMLWVALFGGSRVVHIDPFQKKE-INSIS 231

Qy        249 IPDPQVTSVAFGGPNLDELHVTSAGLQLDD 278

Db        232 VPAKYVTCCAFGGRDLKTLYITTATEQMTE 261


RESULT 9
B90445
regucalcin homolog [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C;Accession: B90445
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: B90445
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-285 <KUR>
A;Cross-references: GB:AE006641; NID:g13816023; PIDN:AAK42817.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO2705
C;Superfamily: senescence marker protein-30

  Query Match             22.6%;  Score 365.5;  DB 2;  Length 285;
  Best Local Similarity   31.6%;  Pred. No. 4.9e-22;
  Matches   98;  Conservative   55;  Mismatches  106;  Indels   51;  Gaps   12;

Qy         13 KYTVGEGPHWDHETQTLYFVD-----------TVEKTFHKYVPSQKKYTFCKVD--KLV 58

Db         12 KATLGEGPVYDKELNKLFWVDIEGKRIIVNDLNTGTEVFYE-MPDLIS-SLCVIDDKRVI 69

Qy         59 SFIIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLG 118

Db         70 ATI-----RHGFYIVNLDR-------------GEIQKIAEVETDMESNRFNDGKCDKLG 110

Qy        119 NLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGK 178

Db        111 RYWAGTMNMNERKP----TGSFYKLDTGKITKILE-GLTVSNGLGWDPDNTQMFLIDSPV 165

Qy        179 RRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQ 238

Db        166 RKVFVFDFDLNKGEIYNRRVAVDFGNE--PGNPDGMAVDEEGYIWVAHWGGGKVSRWNPK 223

Qy        239 QPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGV 298

Db        224 TGKKLFE-IKVPATYVTSVTFGTQELDQLFITTAGKSQDP-------LAGKTF-TTKANV 274

Qy        299 KGFAGVKVKL 308

Db        275 RGLQNFRFKI 284


RESULT 10
A87351
Smp-30/Cgr1 family protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: A87351
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87351
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-289 <STO>
A;Cross-references: GB:AE005673; NID:g13422069; PIDN:AAK22805.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC0820
C;Superfamily: senescence marker protein-30

  Query Match             20.5%;  Score 331;  DB 2;  Length 289;
  Best Local Similarity   33.3%;  Pred. No. 3.1e-19;
```

```
  Matches   94;  Conservative   47;  Mismatches  121;  Indels   20;  Gaps   10;

Qy         13 KYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQ-KKYTFCKVDKLVSFIIPLAGSPGRF 71
              | |:|||| |   :  ||:|||  ::  | | |:  ::::|   |: |:|: |: |: | |
Db         12 KATLGEGPIWHGD--TLWFVDIKQRKIHNYHPATGERFSFDAPDQ-VTFLAPIVGATG-F 67

Qy         72 VVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGL 131
              || |:  |           |    :: ||    ||| ||   |  | || |||
Db         68 VVGLKTGIHRF------HPATGFSLLLEVEDAALNNRPNDATVDAQGRLWFGTMHDGE-- 119

Qy        132 PVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTL 191
                  :|||| :      :| : :| | ||   | | |  |: |: :: :  :|     |
Db        120 --ENNSGSLYRMDLTGVARM-DRDICITNGPCVSPDGKTFYHTDTLEKTIYAFDLAEDGL 176

Qy        192 SISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPD 251
               :||:|    |   :   |||| :| || || |:: |   :: | |   |    :::|
Db        177 -LSNKRVFVQFALGD-DVYPDGSVVDSEGYLWTALWGGFGAVRFSPQGDAV--TRIELPA 232

Qy        252 PQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRV 293
              | ||   ||||:|  |:|:|:|  | | :| :  : | |: |
Db        233 PNVTKPCFGGPDLKTLYFTTARKGLSDETLAQYPLAGGVFAV 274


RESULT 11
A87473
Smp-30/Cgr1 family protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: A87473
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87473
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <STO>
A;Cross-references: GB:AE005673; NID:g13423239; PIDN:AAK23781.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC1805
C;Superfamily: senescence marker protein-30

  Query Match             19.5%;  Score 315;  DB 2;  Length 293;
  Best Local Similarity   30.9%;  Pred. No. 6.2e-18;
  Matches   93;  Conservative   51;  Mismatches  115;  Indels   42;  Gaps   13;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75
              :|| | || :   ||:||::     :| |  :       | :  ::   | ||  :  |
Db         15 LGESPLWDADAGVLYWVDSMAPAIWRYDPFTSEQRSIPAPKPIGSVV--LGRPGELIAGL 72

Qy         76 EREIAILTWDGVSAAPTSI-EAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAI--DAGLP 132
              :  :  |  :  | :: : :  :|      | ||||||  |   |||||:  : |
Db         73 ADGVYRVQLDTGAFTPIALPDTLAPIE------RFNDGKADRQGRFVTGTMAMHNETGR- 125

Qy        133 VGPVTGSLYHLGADKKVKMHESN-IAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTL 191
                 | ||   |    ::  :  | |||   :|     :|: || :  |  : ||  |
Db        126 ----IGKLYRFSAGGAWEVLPTEPIEIANSTCFSPSGDTLYFADSLRHMVWAFSYDPKTG 181

Qy        192 SISNQRPLFTFEKHEVPGY---PDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLD-TV 247
              ::  :|  |      :   |:   ||| |:| ||::|:|: | |::|:||   |:  ||  |
Db        182 AVGEKRDFF-----DTTGFNSAPDGATVDAEGHIWLALVQAQKLIRIS---PDGRLDRVV 233

Qy        248 KIPDPQVTSVAFGGPNLDELHVTSAGLQLDDS------SLDKSLVNGHVYRVTGLGVKGF 301
              : | |  :  |||| :|| |:|||    : ||       :| |   | :    ||||:|
Db        234 ESPAPFCSCPAFGGEDLDILYVTS----ISDSGGRLKTDVDAS---GRLMAFHGLGVRGI 286

Qy        302 A 302
              |
Db        287 A 287
```

```
RESULT 12
G90485
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C;Accession: G90485
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: G90485
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-285 <KUR>
A;Cross-references: GB:AE006641; NID:g13816441; PIDN:AAK43142.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO3041
C;Superfamily: senescence marker protein-30

  Query Match             18.1%;  Score 292;  DB 2;  Length 285;
  Best Local Similarity   30.1%;  Pred. No. 4.3e-16;
  Matches   85;  Conservative   50;  Mismatches  101;  Indels   46;  Gaps   12;

Qy          7 KIAELGKY--TVGEGPHWDHETQTLYFVDTVEK--TFHKYVPSQKKY--------TFCKV 54
              ::  || |   :|| | :| |   |::||   |    :    :::||         : | :
Db          4 ELKTLGNYKAILGESPVYDREGNRLFWVDIDGKKVIVNNLDTNEEKYYDTPDLVTSLCLI 63

Qy         55 DKLVSFIIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIK-----NNRL 109
              |              | :|:|      |          |: |    : |::;      ||
Db         64 DY------------KRIIVTLRHGFHTLNL-------TTGE----ITPYLELENNGENRF 100

Qy        110 NDGKADPLGNLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLK 169
              |||| |     | ||| :    |    | ||| |   :  |: :  :  :||| |  | |
Db        101 NDGKCDIRDRYWAGTMNLHIENP--KPTASLYKLEGQRLTKVLD-KLYRSNGLGWDPDNK 157

Qy        170 KMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQG 229
                | ||:  |:| ::|:|     | |:| :  | |:| || ||| |||||| ||:|   |
Db        158 LFYLIDTPLRKVFQFDFDKDKGDIYNKRIVMDF-KYE-PGRPDGMTIDEEGYLWIAHSAG 215

Qy        230 QRIIKISTQQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTS 271
               :: : : :  | : : :|:|   |:|| || | :::: :|:
Db        216 GKVSRWNPKNGEKVSE-IKLPVLYVSSVTFGSPEMNQIFITT 256


RESULT 13
D95265
probable regucalcin [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSy
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: D95265
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: D95265
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK64686.1; PID:g14523085; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
```

```
A;Gene: SMa0060
A;Genome: plasmid
C;Superfamily: senescence marker protein-30

  Query Match             17.9%;  Score 288.5;  DB 2;  Length 311;
  Best Local Similarity   28.1%;  Pred. No. 9.4e-16;
  Matches   85;  Conservative   47;  Mismatches  128;  Indels   43;  Gaps   11;

Qy         13 KYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAG--SPGR 70
              |  |||   |  : : ||:|| | |  |:  |   ::        |: |    |    |
Db         12 KDIVGESILWCGDEKALYWVDIVGKRIHRLEPENGRHDTWPTPDFVTSI----GMRKDGG 67

Qy         71 FVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMA--ID 128
              |:| | | : : | ||        |     || :  ||||:|:  | |: |  ||   :|
Db         68 FIVGLSRNVCLWTPDG------PFEEFAMPEPDLPENRLNEGRVAPDGSFWVATMQSNLD 121

Qy        129 AG---LPVGPVTGSLYHLGADKKV-KMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEY 184
              ||     :   :|::| :     | ::  :   | | : |: |  : :: |:    :  :
Db        122 AGGSPKDMDRQSGAVYRIDPTGHVSQLTPNEYGITNTMGWTRD-NRFFFADTLANEIYMF 180

Qy        185 DYDASTLSISNQRPLFT-FEKHEVPGYPDGQTIDEEGNLW---------VAVFQGQRIIK 234
              | | :   | |:| :   | :    | |||  :| :  ||          || | |
Db        181 DCDLAARRIDNRRTIVAGFAR----GLPDGSCLDADDRLWNCRVAGGAAVAGFDG----- 231

Qy        235 ISTQQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVT 294
                 |:  :::|    ||  |||| |  |:|||    :    ||   : | :: |
Db        232 -----AGRLMHLIELPASWPTSCTFGGPVLSTLYVTSARFTMTGDHLDMHPLEGGLFAVE 286

Qy        295 GLG 297
              |:|
Db        287 GVG 289


RESULT 14
AB3071
calcium-binding protein, regucalcin [imported] - Agrobacterium tumefaciens (strain C58,
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AB3071
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen; L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB3071
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44984.1; PID:g17742642; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4190
A;Map position: linear chromosome
C;Superfamily: senescence marker protein-30

  Query Match             16.9%;  Score 273;  DB 2;  Length 293;
  Best Local Similarity   31.1%;  Pred. No. 1.6e-14;
  Matches   82;  Conservative   39;  Mismatches  113;  Indels   30;  Gaps   10;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPL-AGSPG----- 69
              ||| | ||  |  |:||| :        ||          ||  | :|   |  |
Db         21 VGESPTWDERTGDLWFVDILAPAIFCLSPS---------GKLQRFDMPAQIGCLGLCQSN 71

Qy         70 RFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDA 129
                | :|:  : :|     :     :| : : :    ::||||||  | |: | ||   :|
Db         72 LIVAALKTGVHLL-----NPQNGELELLCDPDEGRADSRLNDGKVGPDGHFWVGTRD-EA 125

Qy        130 GLPVGPVTGSLYHLGADKKV-KMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
              :| |     || : :   | :  : ::  :||||||  |  :||: ||    :  :|:|
Db        126 DVPTG--NARLYRVSSSGHVERFIDGDMLTSNGLAWSPDETRMYHSDSSGLLLQVFDFDV 183

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPE-VLLDTV 247
                |  :  : |  |:  |  | ||| . | ||  | |  |  |: : :   |:  | :
Db        184 QTGRLGPAKRLHDFQPDE--GRPDGAATDSEGCYWSAGVQAGRLNRFT---PDGELFEIY 238

Qy        248 KIPDPQVTSVAFGGPNLDELHVTS 271
              |:|    |   |||| |  |:|||
Db        239 KMPFKGPTMPCFGGPELKTLYVTS 262


RESULT 15
G98215
senescence marker protein-30 (AB033368) [imported] - Agrobacterium tumefaciens (strai
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 15-Feb-2002
C;Accession: G98215
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: G98215
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-322 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89249.1; PID:g15159075; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_1336
A;Map position: linear chromosome
C;Superfamily: senescence marker protein-30

  Query Match             16.9%;  Score 273;  DB 2;  Length 322;
  Best Local Similarity   31.1%;  Pred. No. 1.8e-14;
  Matches   82;  Conservative   39;  Mismatches  113;  Indels   30;  Gaps   10;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPL-AGSPG----- 69
              ||| | ||  |  |:||| :        ||          ||  | :|   |  |
Db         50 VGESPTWDERTGDLWFVDILAPAIFCLSPS---------GKLQRFDMPAQIGCLGLCQSN 100

Qy         70 RFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDA 129
                | :|:  : :|     :     :| : : :    ::||||||  | |: | ||   :|
Db        101 LIVAALKTGVHLL-----NPQNGELELLCDPDEGRADSRLNDGKVGPDGHFWVGTRD-EA 154

Qy        130 GLPVGPVTGSLYHLGADKKV-KMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
              :| |     || : :   | :  : ::  :||||||  |  :||: ||    :  :|:|
Db        155 DVPTG--NARLYRVSSSGHVERFIDGDMLTSNGLAWSPDETRMYHSDSSGLLLQVFDFDV 212

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPE-VLLDTV 247
                |  :  : |  |:  |  | ||| . | ||  | |  |  |: : :   |:  | :
Db        213 QTGRLGPAKRLHDFQPDE--GRPDGAATDSEGCYWSAGVQAGRLNRFT---PDGELFEIY 267

Qy        248 KIPDPQVTSVAFGGPNLDELHVTS 271
              |:|    |   |||| |  |:|||
Db        268 KMPFKGPTMPCFGGPELKTLYVTS 291


Search completed: August  1, 2003, 14:35:00
Job time : 41 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:25:59 ; Search time 22 Seconds
(without alignments)
580.669 Million cell updates/sec

Title: US-10-089-986-2
Perfect score: 1615
Sequence: 1 MGPVVEKIAELGKYTVGEGP.....HVYRVTGLGVKGFAGVKVKL 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	459	28.4	299	1	SM30_RABIT	Q9ttj6 oryctolagus
2	457.5	28.3	299	1	SM30_MOUSE	Q64374 mus musculu
3	454.5	28.1	299	1	SM30_BOVIN	Q9ttj5 bos taurus
4	454.5	28.1	299	1	SM30_RAT	Q03336 rattus norv
5	448.5	27.8	299	1	SM30_HUMAN	Q15493 homo sapien
6	391.5	24.2	275	1	Y31K_SULAC	P46218 sulfolobus
7	376	23.3	292	1	YVRE_BACSU	O34940 bacillus su
8	235	14.6	340	1	CGR1_CANAL	P56553 candida alb
9	124.5	7.7	320	1	GNL_ZYMMO	Q01578 zymomonas m
10	100	6.2	358	1	YBQ3_YEAST	P38235 saccharomyc
11	94.5	5.9	409	1	EFTU_CHLVU	P56292 chlorella v
12	94	5.8	401	1	PGK_HALVA	P50315 haloarcula
13	92.5	5.7	4351	1	FAT2_RAT	O88277 rattus norv
14	92	5.7	4349	1	FAT2_HUMAN	Q9nyq8 homo sapien
15	91.5	5.7	242	1	RL1_STRVG	P48951 streptomyce
16	91.5	5.7	663	1	ABA2_NICPL	Q40412 nicotiana p
17	91	5.6	396	1	XYNA_BACHD	P07528 bacillus ha
18	91	5.6	568	1	FTSI_BUCAP	O85297 buchnera ap
19	90.5	5.6	242	1	RL1_STRSQ	Q07976 streptomyce
20	90.5	5.6	539	1	CH61_BRAJA	P77829 bradyrhizob
21	90.5	5.6	545	1	CH63_BRAJA	P35862 bradyrhizob
22	89.5	5.5	387	1	XYLA_THETH	P26997 thermus the
23	89	5.5	892	1	LDL2_XENLA	Q99088 xenopus lae
24	88.5	5.5	294	1	OCCT_RHIME	P72298 rhizobium m
25	88.5	5.5	431	1	SPAL_SALTY	P39444 salmonella
26	88	5.4	922	1	DPO1_RICFE	Q9raa9 rickettsia
27	87.5	5.4	242	1	RL1_STRAU	Q9x521 streptomyce
28	87.5	5.4	862	1	LOX1_HORVU	P29114 hordeum vul
29	87	5.4	745	1	RNR_BUCAI	P57628 buchnera ap
30	87	5.4	1071	1	CARB_BACSU	P25994 bacillus su
31	86.5	5.4	720	1	VP19_BPAPS	Q9t1s9 bacteriopha
32	86	5.3	405	1	DHMH_METME	Q59542 methylophil
33	85	5.3	350	1	RLA0_HALVO	P41198 halobacteri
34	85	5.3	426	1	CGED_BACSU	P42092 bacillus su
35	85	5.3	504	1	FLIC_SALMO	Q06973 salmonella
36	85	5.3	636	1	UBA2_YEAST	P52488 saccharomyc
37	85	5.3	992	1	YV93_MYCTU	O53339 mycobacteri
38	85	5.3	1101	1	ACLY_HUMAN	P53396 homo sapien
39	84.5	5.2	541	1	YELL_DROYA	Q9bi17 drosophila
40	84.5	5.2	757	1	AMO_ECOLI	P46883 escherichia
41	84.5	5.2	1139	1	HMW1_MYCGE	Q49413 mycoplasma
42	84	5.2	4660	1	LRP2_RAT	P98158 rattus norv
43	83.5	5.2	538	1	CH60_STAAU	Q08854 staphylococ
44	83.5	5.2	771	1	ITB2_MOUSE	P11835 mus musculu
45	83.5	5.2	935	1	COPG_YEAST	P32074 saccharomyc

ALIGNMENTS

RESULT 1
SM30_RABIT
ID SM30_RABIT STANDARD; PRT; 299 AA.
AC Q9TTJ6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN RGN OR SMP30.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20351777; PubMed=10891565;
RA Misawa H., Yamaguchi M.;
RT "The gene of Ca2+-binding protein regucalcin is highly conserved in vertebrate species.";
RL Int. J. Mol. Med. 6:191-196(2000).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF CALCIUM SIGNALING IN THE AGED LIVER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC ---

CC ---
DR EMBL; AB035445; BAA88079.1; -.
KW Calcium-binding.
SQ SEQUENCE 299 AA; 33108 MW; A8F94AD6596A0C68 CRC64;

Query Match 28.4%; Score 459; DB 1; Length 299;
Best Local Similarity 38.1%; Pred. No. 2.9e-29;
Matches 114; Conservative 41; Mismatches 116; Indels 28; Gaps 8;

```
Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSLE 76
              || | |:  : :| |||   | | :: |  |      :|  |: :       | :| ::
Db         17 GESPVWEEASGSLLFVDIPGKKFCRWNPLTKAVQRMTMDAPVTSV--ALRKSGGYVATVG 74

Qy         77 REIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPVGPV 136
                 :  | :       |: |:  |:   |||| |||| || |  : |||| :      |
Db         75 TKFCALNLE-----DQSVVALATVDKDKKNNRFNDGKVDPAGRYFAGTMAEE----TAPA 125

Qy        137 T-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTL 191
                  |||| |  | :|| :   : |:||| || | |  |||||    || :|||  |
Db        126 VLERHQGSLYALFPDHQVKKYFDQVDISNGLDWSLDHKIFYYIDSLAYSVDAFDYDLQTG 185

Qy        192 SISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDTVKI 249
```

```
Db        186 QISNRRSIYKLEKEE--QIPDGMCIDTEGKLWVACYNGGRVIRL---DPETGKRLQTVKL 240

Qy        250 PDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG-----FAG 303

Db        241 PVDKTTSCCFGGKDYSEMYVTCARDGLDPDSLSRQPEAGGIFKITGLGVKGIPPYSYAG 299


RESULT 2
SM30_MOUSE
ID   SM30_MOUSE     STANDARD;      PRT;   299 AA.
AC   Q64374; Q60944;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN   RGN OR SMP30.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Liver;
RX   MEDLINE=96328264; PubMed=8765750;
RA   Fujita T., Shirasawa T., Maruyama N.;
RT   "Isolation and characterization of genomic and cDNA clones encoding
RT   mouse senescence marker protein-30 (SMP30).";
RL   Biochim. Biophys. Acta 1308:49-57(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=97422495; PubMed=9278263;
RA   Murata T., Yamaguchi M.;
RT   "Molecular cloning of the cDNA coding for regucalcin and its mRNA
RT   expression in mouse liver: the expression is stimulated by calcium
RT   administration.";
RL   Mol. Cell. Biochem. 173:127-133(1997).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY
CC       IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC       CALCIUM SIGNALING IN THE AGED LIVER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: MAINLY PRESENT IN THE LIVER. WEAK EXPRESSION
CC       WAS FOUND IN THE BRAIN AND LUNG, NOT FOUND IN THE KIDNEY.
CC   -!- DEVELOPMENTAL STAGE: PROTEIN AMOUNTS IN LIVER DECREASE
CC       SIGNIFICANTLY WITH AGE.
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28937; AAC52721.1; -.
DR   EMBL; U32170; AAD03478.1; -.
DR   EMBL; D86217; BAA13046.1; -.
DR   SWISS-2DPAGE; Q64374; MOUSE.
DR   MGD; MGI:108024; Rgn.
KW   Calcium-binding.
SQ   SEQUENCE   299 AA;  33407 MW;  DAD55EF618311977 CRC64;

  Query Match           28.3%;  Score 457.5;  DB 1;  Length 299;
  Best Local Similarity   37.8%;  Pred. No. 3.8e-29;
  Matches  112;  Conservative   40;  Mismatches  121;  Indels   23;  Gaps    7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73

Db         14 YRCGESPVWEEASQSLLFVDIPSKIICRWDTVSNQVQRVAVDAPVSSV--ALRQLGGYVA 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133

Db        123 APAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246

Db        183 QTGQISNRRIVYKMEKDE--QIPDGMCIDAEGKLWVACYNGGRVIRL---DPETGKRLQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302

Db        238 VKLPVDKTTSCCFGGKDYSEMYVTCARDGLNAEGLLRQPDAGNIFKITGLGVKGIA 293


RESULT 3
SM30_BOVIN
ID   SM30_BOVIN     STANDARD;      PRT;   299 AA.
AC   Q9TTJ5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN   RGN OR SMP30.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=20351777; PubMed=10891565;
RA   Misawa H., Yamaguchi M.;
RT   "The gene of Ca2+-binding protein regucalcin is highly conserved in
RT   vertebrate species.";
RL   Int. J. Mol. Med. 6:191-196(2000).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY
CC       IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC       CALCIUM SIGNALING IN THE AGED LIVER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB035446; BAA88080.1; -.
KW   Calcium-binding.
SQ   SEQUENCE   299 AA;  33308 MW;  4258F7CDB3CCB575 CRC64;

  Query Match           28.1%;  Score 454.5;  DB 1;  Length 299;
  Best Local Similarity   38.1%;  Pred. No. 6.6e-29;
  Matches  111;  Conservative   38;  Mismatches  119;  Indels   23;  Gaps    7;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSLE 76

Db         17 GESPVWEEASNSLLFVDIPAKKVCRWDSLSKQVQRVTVDAPVSSV--ALRQSGGYVATVG 74

Qy         77 REIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPVGPV 136

Db         75 TKFCALNWEDQSAV-----VLATVDKEKKNNRFNDGKVDPAGRYFAGTMAEE----TAPA 125

Qy        137 T-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTL 191

Db        126 VLERRQGSLYSLFPDHHVEKYFDQVDISNGLDWSMDHKIFYYIDSLSYSVDAFDYDLQTG 185

Qy        192 SISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDTVKI 249
```

```
                ||||:| ::  || |    |||  || || |||| : | |:|::    ||    | |||:
Db        186 KISNRRSVYKLEKEE--QIPDGMCIDVEGKLWVACYNGGRVIRL---DPETGKRLQTVKL 240

Qy        250 PDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG 300
                |  : ||  ||| :  |::|| |   ||   | :    | ::::|||||||
Db        241 PVDKTTSCCFGGKDYSEMYVTCARDGLDSKGLLQQPEAGGIFKITGLGVKG 291


RESULT 4
SM30_RAT
ID   SM30_RAT       STANDARD;      PRT;   299 AA.
AC   Q03336; Q63496;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN   RGN OR SMP30.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=93041931; PubMed=1420310;
RA   Fujita T., Shirasawa T., Uchida K., Maruyama N.;
RT   "Isolation of cDNA clone encoding rat senescence marker protein-30
RT   (SMP30) and its tissue distribution.";
RL   Biochim. Biophys. Acta 1132:297-305(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=93351639; PubMed=8348951;
RA   Shimokawa N., Yamaguchi M.;
RT   "Molecular cloning and sequencing of the cDNA coding for a calcium-
RT   binding protein regucalcin from rat liver.";
RL   FEBS Lett. 327:251-255(1993).
RN   [3]
RP   SEQUENCE OF 117-299 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX   MEDLINE=96150990; PubMed=8569761;
RA   Shimokawa N., Matsuda Y., Yamaguchi M.;
RT   "Genomic cloning and chromosomal assignment of rat regucalcin gene.";
RL   Mol. Cell. Biochem. 151:157-163(1995).
RN   [4]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=96386712; PubMed=8794449;
RA   Fujita T., Shirasawa T., Uchida K., Maruyama N.;
RT   "Gene regulation of senescence marker protein-30 (SMP30): coordinated
RT   up-regulation with tissue maturation and gradual down-regulation
RT   with aging.";
RL   Mech. Ageing Dev. 87:219-229(1996).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY
CC       IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC       CALCIUM SIGNALING IN THE AGED LIVER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: HEPATOCYTES AND RENAL PROXIMAL TUBULAR
CC       EPITHELIUM.
CC   -!- DEVELOPMENTAL STAGE: IN LIVER, THE FIRST PEAK OF EXPRESSION WAS
CC       FOUND IN 5-DAY-OLD NEONATES. EXPRESSION INCREASES FROM DAY 7 AND
CC       REACHES A PLATEU AT DAY 10. 3-6.5 MOTH-OLD ADULTS EXPRESS ABOUT A
CC       THIRD THE AMOUNT OF NEONATES LEVEL. IN KIDNEY, EXPRESSION
CC       INCREASES FROM DAY 21 AND REACHES A MAXIMAL LEVEL AT DAY 35,
CC       REMAINS HIGH UNTIL 3 MONTHS OF AGE.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69021; CAA48786.1; -.
DR   EMBL; D38467; BAA07490.1; -.
DR   EMBL; D31662; BAA06507.1; -.
DR   PIR; S27203; S27203.
KW   Calcium-binding.
FT   CONFLICT    148    148       N -> D (IN REF. 1).
SQ   SEQUENCE   299 AA;  33389 MW;  E3CF6D3FFCAE4E98 CRC64;

  Query Match             28.1%;  Score 454.5;  DB 1;  Length 299;
  Best Local Similarity   37.5%;  Pred. No. 6.6e-29;
  Matches  111;  Conservative   41;  Mismatches  121;  Indels   23;  Gaps    7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              |  || | |:  :: | |||   ||  ::     :     ||  || :       | :|
Db         14 YRCGESPVWEEASKCLLFVDIPSKTVCRWDSISNRVQRVGVDAPVSSV--ALRQSGGYVA 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::  :   | |:          |:  :  |:   |||| |||| || |  : |||| :
Db         72 TIGTKFCALNWE-----DQSVFILAMVDEDKKNNRFNDGKVDPAGRYFAGTMAEE----T 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
                |       |||| |  |  || : : : |:||| || | |  |||||    || :|||
Db        123 APAVLERHQGSLYSLFPDHSVKKYFNQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
                | |||:| ::  || |    |||  || || |||| : | |:|::    ||    | |
Db        183 PTGQISNRRTVYKMEKDE--QIPDGMCIDVEGKLWVACYNGGRVIRL---DPETGKRLQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||| :  |::|| |   :    | :    |:::::|||||| |
Db        238 VKLPVDKTTSCCFGGKDYSEMYVTCARDGMSAEGLLRQPDAGNIFKITGLGVKGIA 293


RESULT 5
SM30_HUMAN
ID   SM30_HUMAN     STANDARD;      PRT;   299 AA.
AC   Q15493;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
GN   RGN OR SMP30.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96004897; PubMed=7548213;
RA   Fujita T., Mandel J.L., Shirasawa T., Hino O., Shirai T., Maruyama N.;
RT   "Isolation of cDNA clone encoding human homologue of senescence
RT   marker protein-30 (SMP30) and its location on the X chromosome.";
RL   Biochim. Biophys. Acta 1263:249-252(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   PubMed=10677570;
RA   Misawa H., Yamaguchi M.;
RT   "Transcript heterogeneity of the human gene for Ca2+-binding protein
RT   regucalcin.";
RL   Int. J. Mol. Med. 5:283-287(2000).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY
CC       IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
CC       CALCIUM SIGNALING IN THE AGED LIVER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D31815; BAA06602.1; -.
DR   EMBL; AB028125; BAA78693.1; -.
DR   EMBL; AB032064; BAA84082.1; -.
DR   Genew; HGNC:9989; RGN.
DR   MIM; 300212; -.
KW   Calcium-binding.
SQ   SEQUENCE   299 AA;  33253 MW;  95BA1C73B7B77635 CRC64;

  Query Match           27.8%;  Score 448.5;  DB 1;  Length 299;
  Best Local Similarity  36.9%;  Pred. No. 2e-28;
  Matches  110;  Conservative   42;  Mismatches  113;  Indels   33;  Gaps    8;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKY---TFCKVDKLVSFIIPLAG----SPG 69
              || | |:  : :| |||       .  :|::|     :| |  : |:   |::       |
Db         17 GESPVWEEVSNSLLFVD---------IPAKKVCRWDSFTKQVQRVTMDAPVSSVALRQSG 67

Qy         70 RFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDA 129
                :| ::  :   | |   ||        :  |:   ||||.|||| || |  : |||| :
Db         68 GYVATIGTKFCALNWKEQSAV-----VLATVDNDKKNNRFNDGKVDPAGRYFAGTMAEE- 121

Qy        130 GLPVGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEY 184
                 |       |:|| |  |  || :   : |:||| || | |  |||||    || :
Db        122 ---TAPAVLERHQGALYSLFPDHHVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYSVDAF 178

Qy        185 DYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLL 244
              |||  |  |||:| ::  || |    |||  || || |||| : | |:|::     :  |
Db        179 DYDLQTGQISNRRSVYKLEKEE--QIPDGMCIDAEGKLWVACYNGGRVIRLDPVTGK-RL 235

Qy        245 DTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
               |||:|  : ||  ||| |  |::|| |   :|   | :   | :::|||||||| |
Db        236 QTVKLPVDKTTSCCFGGKNYSEMYVTCARDGMDPEGLLRQPEAGGIFKITGLGVKGIA 293


RESULT 6
Y31K_SULAC
ID   Y31K_SULAC     STANDARD;      PRT;   275 AA.
AC   P46218;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Hypothetical 31.5 kDa protein.
OS   Sulfolobus acidocaldarius.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=2285;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Durovic P.V., Potter S., Dennis P.P.;
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE A RNA POLYMERASE
CC       SUBUNIT.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U05664; AAA73413.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   275 AA;  31544 MW;  D949D5BF83EBE582 CRC64;

  Query Match           24.2%;  Score 391.5;  DB 1;  Length 275;
  Best Local Similarity  33.0%;  Pred. No. 5.7e-24;
  Matches   98;  Conservative   51;  Mismatches   97;  Indels   51;  Gaps   10;

Qy         18 EGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIP-------LAGSPGR 70
              ||| | :   :|||||  :   | :     :   |    || : |       :    |
Db         13 EGPIWAY--NSLYFVDIPKGELHN-LKEDGTHWVVKFPTYVSSLQPTKRGGIIVTAGNGF 69

Qy         71 FVVSLEREIAIL----TWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMA 126
              ::|  : :|::|     ||                        || |||| | :|  | |||
Db         70 YLVKDKDQISLLYEVKDWD-----------------SRNRFNDGKCDQMGRYWIGTMN 110

Qy        127 IDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDY 186
              ::   |    || |: |  | :   ::: |:|||||| | | :|||||  |:: :: :
Db        111 LEEKYP----TGGLFVLDLDMKFRRVLTDVTISNGLAWSLDNKYLYYIDSPTRKIFKFKF 166

Qy        187 DASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDT 246
              |    || :  |   :::|  | ||| ||| |||||||:: |  :::|  :: :|: :
Db        167 DLERGDISQREVLIDLKEYE--GVPDGMTIDSEGNLWVALYGGGAVLRIDVEKRKVIQE- 223

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVY--RVTGLGVKGF 301
              :::| |:|||| ||| |:| | :|:|    |          | ||  ||   ||: :
Db        224 LRLPAPRVTSVIFGGSNMDTLFITTANDHPD---------GGFVYSERVDVKGVETY 271


RESULT 7
YVRE_BACSU
ID   YVRE_BACSU     STANDARD;      PRT;   292 AA.
AC   O34940;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical protein yvrE.
GN   YVRE.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98304083; PubMed=9639930;
RA   Wipat A., Brignell C.S., Guy J.B., Rose M., Emmerson P.T.,
RA   Harwood C.R.;
RT   "The yvsA-yvqA (293 degrees - 289 degrees) region of the Bacillus
RT   subtilis chromosome containing genes involved in metal ion uptake and
RT   a putative sigma factor.";
RL   Microbiology 144:1593-1600(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
```

```
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ223978; CAA11732.1; -.
DR   EMBL; Z99120; CAB15310.1; -.
DR   EMBL; Z99121; CAB15325.1; -.
DR   SubtiList; BG14145; yvrE.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   292 AA;  33204 MW;  D68A0812FC8D90AB CRC64;

  Query Match             23.3%;  Score 376;  DB 1;  Length 292;
  Best Local Similarity   31.5%;  Pred. No. 1e-22;
  Matches   85;  Conservative   52;  Mismatches  105;  Indels   28;  Gaps    7;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75
              :|||| || |   ||:|| :    | : | :      |::: : |          || :|
Db         13 IGEGPLWDEENGRLYWVDILGSELHIFDPEE------KINRSIKF--------KSFVTAL 58

Qy         76 ER----EIAILTWDG---VSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAID 128
                :   |: :   ||         |:| |   :   :: | || | || | || || ::: :
Db         59 AKYSKDELIMTMKDGFYLYHLRDDSLEKIKQPKDMHESLRFNDAKCDPYGRLWAGTTSME 118

Qy        129 AGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
                  ||| |  |  :   :  :: :||| |  :   |||||:  : :  | ||
Db        119 G----EQKQASLYRLNLDGSLVKIKDQVSTSNGLDWDRERNLMYYIDTPTQEIVRYSYDP 174

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVK 248
               :  :||  |:: |:: :  | ||| |||: | ||||:| | |:: |  _| :  ::::
Db        175 QSGDVSNPEPVYRFDQSD--GLPDGMTIDQNGMLWVALFGGSRVVHIDPFQKKE-INSIS 231

Qy        249 IPDPQVTSVAFGGPNLDELHVTSAGLQLDD 278
              :|   ||  |||| :|  |::|:|  |: :
Db        232 VPAKYVTCCAFGGRDLKTLYITTATEQMTE 261


RESULT 8
CGR1_CANAL
ID   CGR1_CANAL     STANDARD;      PRT;   340 AA.
AC   P56553;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Cell growth protein CGR1.
GN   CGR1.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 9061;
RA   Cho T., Sudoh M., Tanaka T., Nakashima Y.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE CELL GROWTH REGULATION.
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB013716; BAA28176.1; -.
SQ   SEQUENCE   340 AA;  38391 MW;  4B1673ED448D1F7A CRC64;

  Query Match             14.6%;  Score 235;  DB 1;  Length 340;
  Best Local Similarity   26.9%;  Pred. No. 1.9e-11;
  Matches   84;  Conservative   55;  Mismatches  123;  Indels   50;  Gaps   15;

Qy         12 GKYTVGEGPHWDHETQTLYFVDTVEKTFHKYV---------PSQKKYTFCKVDKLVSFI- 61
              |: |  ||  :|    || :|| ::   |:           |   :   | |     |
Db         17 GRLT--EGVTYDERNNTLLWVDIIQGEVHRVFLDNTNTNINTSSSSHETLKWDSSNESIG 74

Qy         62 -IPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNN------RLNDGKA 114
               | |   | : ::  :  :|   :     : ::||      ||  |       | |||
Db         75 AICLTNDPNKLIICSKYGLAYGDF-----SSSTIEYFFKY-PHTTNPDEKLRLRSNDGII 128

Qy        115 DPLGNLWTGTMAIDAGLPVG-----PVTGSLYHLGADKKVK------MHESNIAIANGLA 163
              || |||| | |      |:|        | || :|, |:        | |::: |:|||
Db        129 DPWGNLWIGVM---NDFPIGAKEGIQPEGKLYRIGFSKESNKLTCDVMIENSL-ISNGLC 184

Qy        164 WSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVP----GYPDGQTIDEE 219
              ::|   : |: ||   :: :|:|| :|  ::| : :|   |   |       ||| :
Db        185 FNNQGDEFYWTDSLTFKIWKYNYDKTTNKLTN-KSVFIDLKQFYPDVEQSEPDGLVMTNN 243

Qy        220 GNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDPQVTSVAFG-GPNLDELHVTSAGLQLDD 278
              | ::  ||    |: :  |  |:  : :||   : | |  | |   :|: ||:  |:|||
Db        244 GEIYTCVFSTGTILHVDNQGKEI--ERIKIAAKRPTCVTIGSGIKNNEMFVTTGHLKLDD 301

Qy        279 --SSLDKSLVNG 288
                :::| : ::|
Db        302 EKATIDATNLDG 313


RESULT 9
GNL_ZYMMO
ID   GNL_ZYMMO      STANDARD;      PRT;   320 AA.
AC   Q01578;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   Gluconolactonase precursor (EC 3.1.1.17) (D-glucono-delta-lactone
DE   lactonohydrolase).
GN   GNL.
OS   Zymomonas mobilis.
OC   Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC   Zymomonas.
OX   NCBI_TaxID=542;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 36-58.
RC   STRAIN=ATCC 29191 / ZM6;
RX   MEDLINE=93129619; PubMed=1482681;
RA   Kanagasundaram V., Scopes R.;
RT   "Isolation and characterization of the gene encoding gluconolactonase
RT   from Zymomonas mobilis.";
RL   Biochim. Biophys. Acta 1171:198-200(1992).
CC   -!- FUNCTION: HYDROLYZES THE GLUCONOLACTONE FORMED BY GLUCOSE-FRUCTOSE
CC       OXIDOREDUCTASE, AND THAT FORMED IN AEROBIC CONDITIONS BY THE
CC       GLUCOSE DEHYDROGENASE PRESENT.
CC   -!- CATALYTIC ACTIVITY: D-glucono-1,5-lactone + H(2)O = D-gluconate.
CC   -!- PATHWAY: SECOND STEP IN SORBITOL PATHWAY.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC   -!- CAUTION: IT IS POSSIBLE THAT IN SOME CASES MET-40 IS THE INITIATOR
CC       IN THIS CASE THE EXPRESSED ENZYME WOULD REMAIN CYTOPLASMIC.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X67189; CAA47637.1; -.
DR   PIR; S25124; S25124.
DR   PIR; S28218; S28218.
DR   InterPro; IPR000033; Ldl_receptor_rep.
DR   SMART; SM00135; LY; 1.
KW   Hydrolase; Serine esterase; Periplasmic; Signal.
FT   SIGNAL        1     35
FT   CHAIN        36    320       GLUCONOLACTONASE.
FT   CHAIN        41    320       GLUCONOLACTONASE (MINOR FORM).
SQ   SEQUENCE   320 AA;  34767 MW;  7552DE348D83564C CRC64;

  Query Match             7.7%;  Score 124.5;  DB 1;  Length 320;
  Best Local Similarity   21.8%;  Pred. No. 0.0099;
  Matches   58;  Conservative   26;  Mismatches   83;  Indels   99;  Gaps    9;

Qy         18 EGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSLER 77
              ||| |      | | |       |: |          |  ||  :  :|         |
Db         76 EGPVWVKNGNFLLFSDPPANIMRKWTP---------DAGVSIFLKPSGHAEPIPAGQFR 125

Qy         78 EIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLW---TGTMAIDAGLPVG 134
              |            | |                 |  |  | | :|   :|| ||   : |
Db        126 E------------PGS---------------NGMKVGPDGKIWVADSGTRAI---MKVD 154

Qy        135 PVT-------------------------GSLY-------------------------HLG 144
              |||                        |::|                         |
Db        155 PVTRQRSVVDNYKGKRFNSPNDLFFSKSGAVYFTDPPYGLTNLDESDIKEMNYNGVFRLS 214

Qy        145 ADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLF-TFE 203
               | :: : |: ::  |||| | |  |:|  :| :   : : |   :   :   | |   |
Db        215 PDGRLDLIEAGLSRPNGLALSPDETKLYVSNSDRASPNIWVYSLDSNGLPTSRTLLRNFR 274

Qy        204 K----HEVPGYPDGQTIDEEGNLWVA 225
              |      : | |||  ||::|||: :
Db        275 KEYFDQGLAGLPDGMNIDKQGNLFAS 300


RESULT 10
YBQ3_YEAST
ID   YBQ3_YEAST     STANDARD;      PRT;   358 AA.
AC   P38235;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 40.3 kDa protein in REG2-YRO2 intergenic region.
GN   YBR053C OR YBR0506.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=95321020; PubMed=7597852;
RA   Aljinovic G., Pohl T.M.;
RT   "Sequence and analysis of 24 kb on chromosome II of Saccharomyces
RT   cerevisiae.";
RL   Yeast 11:475-479(1995).
CC   -!- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z35922; CAA84996.1; -.
DR   EMBL; Z46260; CAA86396.1; -.
DR   PIR; S45911; S45911.
DR   SGD; S0000257; YBR053C.
KW   Hypothetical protein.
SQ   SEQUENCE   358 AA;  40295 MW;  0091701F98ED87C9 CRC64;

  Query Match             6.2%;  Score 100;  DB 1;  Length 358;
  Best Local Similarity   24.5%;  Pred. No. 1;
  Matches   52;  Conservative   41;  Mismatches   73;  Indels   46;  Gaps   15;

Qy        108 RLNDGKADPLGN-LWTGTMAIDAGLPVGPVTGSLYHLG-ADKKVKMHESNIAIANGLAW- 164
              | |||   | |  :: | |: |    : |: | |  :     |:::  : : | | : |
Db        148 RSNDGNVSPDGKYIYVGLMS-DFPFDLEPI-GCLLRVDLLAHKIELVWNCLLIPNAIHWD 205

Qy        165 SNDLKKMYYIDSGK-----------RRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPD 212
              :| | ||  ||              :| :  |   |    :||   :|| |   ||
Db        206 ESDQKTMYVTDSLNFTIWKCPGGDLLKRDELIDVKNS----NNQ----SFESPE----PD 253

Qy        213 GQTI------DEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPD--PQVTSVAFGGPNL 264
              |  |         | |:: |:   ::       ::| : : :|:  |:|:   | |
Db        254 GSAIWFSKDGKHSGFLFITVWSTSKVQMFDLTNGKLLKEFI-LPEQTPRVSCCCFVG--- 309

Qy        265 DELHVTSAGLQLDD---SSLDKSLVNGHVYRV 293
               :| ||:|  :::|   :: ||:   | :|::
Db        310 KDLFVTTANAEINDAVRTNTDKN--GGCIYKI 339


RESULT 11
EFTU_CHLVU
ID   EFTU_CHLVU     STANDARD;      PRT;   409 AA.
AC   P56292;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Elongation factor Tu (EF-Tu).
GN   TUFA.
OS   Chlorella vulgaris.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC   Chlorellaceae; Chlorella.
OX   NCBI_TaxID=3077;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IAM C-27 / Tamiya;
RX   MEDLINE=97303241; PubMed=9159184;
RA   Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA   Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA   Inamura A., Yoshinaga K., Sugiura M.;
RT   "Complete nucleotide sequence of the chloroplast genome from the
RT   green alga Chlorella vulgaris: the existence of genes possibly
RT   involved in chloroplast division.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC   -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC       AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC       BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC       EF-TU/EF-1A SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB001684; BAA57886.1; -.
DR   HSSP; P07157; 1AIP.
DR   InterPro; IPR004541; EF-Tu.
DR   InterPro; IPR004160; EFTU_Cterm.
DR   InterPro; IPR004161; EFTU_D2.
DR   InterPro; IPR000795; EF_GTPbind.
DR   InterPro; IPR005225; Small_GTP.
```

```
DR   Pfam; PF00009; GTP_EFTU; 1.
DR   Pfam; PF03143; GTP_EFTU_D3; 1.
DR   Pfam; PF03144; GTP_EFTU_D2; 1.
DR   PRINTS; PR00315; ELONGATNFCT.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   TIGRFAMs; TIGR00485; EF-Tu; 1.
DR   PROSITE; PS00301; EFACTOR_GTP; 1.
KW   Elongation factor; Protein biosynthesis; Chloroplast;
KW   GTP-binding.
FT   NP_BIND      19     26       GTP (BY SIMILARITY).
FT   NP_BIND      81     85       GTP (BY SIMILARITY).
FT   NP_BIND     136    139       GTP (BY SIMILARITY).
SQ   SEQUENCE   409 AA;  44911 MW;  0183143E5E117999 CRC64;

  Query Match             5.9%;  Score 94.5;  DB 1;  Length 409;
  Best Local Similarity   19.6%;  Pred. No. 3.2;
  Matches   70;  Conservative   49;  Mismatches   98;  Indels  141;  Gaps   19;

Qy         15 TVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVS 74
              |:|   | ||   ||    |:                          :    |: ||
Db         17 TIG---HVDHGKTTLTAAITMA-----------------------LAARGGAKGR---- 45

Qy         75 LEREIAILTWDGVSAAPTSIEAIVNVE-PHIKNNRLN--------DGKADPLGNLWTGTM 125
                     :| : :||      : :   |::    |          | || : |: ||
Db         46 --------KYDDIDSAPEEKARGITINTAHVEYETENRHYAHVDCPGHADYVKNMITGAA 97

Qy        126 AIDAGLPV-----GPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRR 180
              :| |: |     ||:  :  ||   |:|        : | : : |             :
Db         98 QMDGGILVVSGADGPMPQTKEHLLLAKQV-------GVPNIVVFLN------------K 137

Qy        181 VDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKIS--TQ 238
              |: | ||  | :        | :|:| ||       ||     : :  |  :: :   |:
Db        138 EDQVD-DAELLELLELEIRETLDKYEFPG-------DE-----IPIISGSALLALEALTE 184

Qy        239 QPEV------------LLDTVK--IPDPQ----------------VT---SVAFG---- 260
              ||              |:| |   || |:               :|   :|| |
Db        185 NPETKPGDNKWVDKIYNLMDQVDSYIPTPERETEKPFLMAVEDVFSITGRGTVATGRVER 244

Qy        261 -----GPNL------DELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVK 307
                  |  :      |    |  ||::   :||:|:   :|    |: ::|   : ::
Db        245 GCVKIGDTVELVGLRDTKTTTVTGLEMFQKTLDESVAGDNV----GILLRGVQKIDIE 298


RESULT 12
PGK_HALVA
ID   PGK_HALVA      STANDARD;      PRT;   401 AA.
AC   P50315;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Phosphoglycerate kinase (EC 2.7.2.3).
GN   PGK.
OS   Haloarcula vallismortis.
OC   Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC   Halobacteriaceae; Haloarcula.
OX   NCBI_TaxID=28442;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29715;
RX   MEDLINE=96197399; PubMed=8616244;
RA   Brinkmann H., Martin W.;
RT   "Higher-plant chloroplast and cytosolic 3-phosphoglycerate kinases: a
RT   case of endosymbiotic gene replacement.";
RL   Plant Mol. Biol. 30:65-75(1996).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC       phospho-D-glyceroyl phosphate.
CC   -!- PATHWAY: Second phase of glycolysis; second step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L47295; AAB03731.1; -.
DR   HSSP; P18912; 1PHP.
DR   InterPro; IPR001576; PGK.
DR   Pfam; PF00162; PGK; 1.
DR   PRINTS; PR00477; PHGLYCKINASE.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW   Transferase; Kinase; Glycolysis.
SQ   SEQUENCE   401 AA;  44193 MW;  ACDD98F8856B490E CRC64;

  Query Match             5.8%;  Score 94;  DB 1;  Length 401;
  Best Local Similarity   22.1%;  Pred. No. 3.5;
  Matches   78;  Conservative   42;  Mismatches  137;  Indels   96;  Gaps   18;

Qy          2 GP-VVEKIAELGKYTV----------GEGPHWDHETQT-LYFVDTVEKTFHKYVPSQKKY 49
              ||  :  ||:|    |           | |  | | :    || |:   |  |:
Db         94 GPQAIHDIADLDSGDVLVLENTRMCDDELPEEDPEVKAQTEFVKTLAGEFDAYINDAYSA 153

Qy         50 TFCKVDKLVSFIIPLAGSPGRFVVSLERE----IAILTWDG---VSAAPTSIEAIVNVEP 102
                    || | : :    || |:  | |    ||   :||   :    |    :::|
Db        154 AHRSHASLVGFPLVMDAYAGR-VMETEYEANTAIAEKEFDGQVTMVVGGTKATDVIDVMT 212

Qy        103 HIKNNRLNDGKADP--LGNLWTGTMAIDAGLPVG------PVTGSLYHLGADKKVKM--- 151
              |:       | | |   || :  ||:   || |||      :   :   ::|   |
Db        213 HL------DEKVDDFLLGGI-AGTVPAAAGHPVGYDIDDANLYDEQWEANSEKIESMLED 265

Qy        152 HESNIAIANGLAWSN-------------DLKKMYYIDSGKRRVDEYDYDASTLSISNQRP 198
              |   | :|  ||: :             | |:: |:| |   : ||
Db        266 HRDQITLAVDLAYEDENDDRAEQAVDDIDEKRLSYLDVGSETLMEYS------------- 312

Qy        199 LFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDPQVTSVA 258
                    |   : : :  ||     :|: :|    :    | : ||         ||
Db        313 ---------PIIRESEAVFGEGR--AGMFEDERFSVGTAGVLEAIADT------DCFSVV 355

Qy        259 FGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG-----FAGVKV 306
              ||        ||  :::     |:   |||    |  ::       ||:|
Db        356 GGGD-------TSRAIEMYGMEEDEF---GHVSIAGGAYIRALTRAQLVGVEV 398


RESULT 13
FAT2_RAT
ID   FAT2_RAT       STANDARD;      PRT;  4351 AA.
AC   O88277;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Protocadherin Fat 2 precursor (Multiple epidermal growth factor-like
DE   domains 1).
GN   FAT2 OR MEGF1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=98360089; PubMed=9693030;
RA   Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT   "Identification of high-molecular-weight proteins with multiple
RT   EGF-like motifs by motif-trap screening.";
RL   Genomics 51:27-34(1998).
CC   -!- SIMILARITY: CONTAINS 33 CADHERIN DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB011527; BAA32458.1; -.
DR   HSSP; P08709; 1BF9.
DR   InterPro; IPR002126; Cadherin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001791; Laminin_G.
DR   Pfam; PF00028; cadherin; 33.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00054; laminin_G; 1.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 33.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00282; LamG; 1.
DR   PROSITE; PS00232; CADHERIN_1; 14.
DR   PROSITE; PS50268; CADHERIN_2; 33.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 2.
KW   Transmembrane; Glycoprotein; Repeat; EGF-like domain; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   4351       PROTOCADHERIN FAT 2.
FT   DOMAIN       19   4050       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   4051   4071       POTENTIAL.
FT   DOMAIN     4072   4351       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       34    148       CADHERIN 1.
FT   DOMAIN      149    256       CADHERIN 2.
FT   DOMAIN      363    458       CADHERIN 3.
FT   DOMAIN      459    564       CADHERIN 4.
FT   DOMAIN      565    669       CADHERIN 5.
FT   DOMAIN      716    820       CADHERIN 6.
FT   DOMAIN      821    925       CADHERIN 7.
FT   DOMAIN      926   1032       CADHERIN 8.
FT   DOMAIN     1033   1142       CADHERIN 9.
FT   DOMAIN     1138   1242       CADHERIN 10.
FT   DOMAIN     1243   1346       CADHERIN 11.
FT   DOMAIN     1350   1448       CADHERIN 12.
FT   DOMAIN     1449   1555       CADHERIN 13.
FT   DOMAIN     1556   1660       CADHERIN 14.
FT   DOMAIN     1661   1758       CADHERIN 15.
FT   DOMAIN     1759   1872       CADHERIN 16.
FT   DOMAIN     1873   1968       CADHERIN 17.
FT   DOMAIN     1969   2070       CADHERIN 18.
FT   DOMAIN     2071   2171       CADHERIN 19.
FT   DOMAIN     2172   2272       CADHERIN 20.
FT   DOMAIN     2273   2379       CADHERIN 21.
FT   DOMAIN     2380   2481       CADHERIN 22.
FT   DOMAIN     2482   2585       CADHERIN 23.
FT   DOMAIN     2586   2692       CADHERIN 24.
FT   DOMAIN     2693   2799       CADHERIN 25.
FT   DOMAIN     2800   2908       CADHERIN 26.
FT   DOMAIN     2909   3013       CADHERIN 27.
FT   DOMAIN     3014   3115       CADHERIN 28.
FT   DOMAIN     3116   3220       CADHERIN 29.
FT   DOMAIN     3221   3323       CADHERIN 30.
FT   DOMAIN     3324   3428       CADHERIN 31.
FT   DOMAIN     3429   3533       CADHERIN 32.
FT   DOMAIN     3534   3631       CADHERIN 33.
FT   DOMAIN     3775   3946       LAMININ G-LIKE.
FT   DOMAIN     3952   3986       EGF-LIKE 1.
FT   DOMAIN     3991   4024       EGF-LIKE 2.
FT   CARBOHYD     39     39       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    210    210       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    568    568       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    627    627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    789    789       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    996    996       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1175   1175       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1276   1276       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1417   1417       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1899   1899       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1998   1998       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2007   2007       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2102   2102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2165   2165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2183   2183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2325   2325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2368   2368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2387   2387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2430   2430       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2470   2470       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2547   2547       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2597   2597       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3127   3127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3278   3278       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3312   3312       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3432   3432       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3603   3603       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3770   3770       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3774   3774       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3815   3815       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3842   3842       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3875   3875       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3906   3906       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3991   3991       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   4351 AA;  480644 MW;  C564E7F98EBB3888 CRC64;

  Query Match             5.7%;  Score 92.5;  DB 1;  Length 4351;
  Best Local Similarity   20.9%;  Pred. No. 1e+02;
  Matches   76;  Conservative   54;  Mismatches  133;  Indels  101;  Gaps   17;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCK----------------------V 54
              |  ||||         |   |    |:  |:  |||                      :
Db       2667 GGPPHWDSLVPVRLQVVPNEIPLPKF--SEPLYTFSAPEDLPEGSEIGSVKAVAAQDPII 2724

Qy         55 DKLVSFIIPLAGSPGRFVVSLEREIAIL----------TW---DGVSAAP---TSIEAI 97
               ||    | : |    | ||::: :|           |   | :: |   | : ::
Db       2725 YSLVQGTTPESNSDD--VFSLDQDTGVLKVRKAMDHESTKWYQIDLMAHCPHEDTDLVSL 2782

Qy         98 VNVEPHIKNNRLNDGK----ADPLGNLWTGTMAIDAGLPVGPVTGSLYHLGADKKV---- 149
              |:|   :::  :|| :    |||     |  |    |  |  || :    |:| :|
Db       2783 VSVSIQVED--VNDNRPVFEADPYKAFLTENM--PGGTTVIQVTANDQDTGSDGQVSYRL 2838

Qy        150 ------KMHESNIAIANGLAWSNDLKKM---------YYIDSGKRRVDEYDYDASTLSIS 194
                   :||   |: :   |   |:::         :|: :       :|:   |: :|
Db       2839 SVEPGSNIHEL-FAVDSESGWITTLQELDCETQQTYRFYVVA-------FDH-GQTIQLS 2889

Qy        195 NQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDP-- 252
              :|  :      |    |   : |  |::           :   :|  |: |:|  |
Db       2890 SQALVEVSITDENDNPPRFASEDYRGSV-----------VENNEPGELVATLKTLDADV 2937

Qy        253 -----QVTSVAFGGPNLDELHVTSAGLQLDDSS---LDKSLVNGHVYRVTGLGVKGFAGV 304
                  |||    |  | :  ::  | :   ||   ||:  :  :: |||    |  | |
Db       2938 SDQNRQVTCYITEGDPLGQFSISQVGDEWRISSRKTLDREHIAKYLLRVTASDGKFQASV 2997

Qy        305 KVKL 308
                |::
Db       2998 PVEV 3001


RESULT 14
FAT2_HUMAN
ID   FAT2_HUMAN     STANDARD;      PRT;  4349 AA.
AC   Q9NYQ8; O75091; Q9NSR7;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Protocadherin Fat 2 precursor (hFat2) (Multiple epidermal growth
DE   factor-like domains 1).
GN   FAT2 OR MEGF1 OR CDHF8.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20202599; PubMed=10716726;
RA   Wu Q., Maniatis T.;
RT   "Large exons encoding multiple ectodomains are a characteristic
RT   feature of protocadherin genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).
RN   [2]
RP   SEQUENCE OF 3777-4349 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98360089; PubMed=9693030;
RA   Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT   "Identification of high-molecular-weight proteins with multiple
RT   EGF-like motifs by motif-trap screening.";
RL   Genomics 51:27-34(1998).
RN   [3]
RP   SEQUENCE OF 4142-4349 FROM N.A.
RC   TISSUE=Testis;
RA   Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 32 CADHERIN DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF231022; AAF61928.1; -.
DR   EMBL; AB011535; BAA32463.1; -.
DR   EMBL; AL157443; CAB75663.1; -.
DR   Genew; HGNC:3596; FAT2.
DR   MIM; 604269; -.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR002126; Cadherin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001791; Laminin_G.
DR   Pfam; PF00028; cadherin; 33.
DR   Pfam; PF00008; EGF; 2.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 32.
DR   SMART; SM00181; EGF; 2.
DR   SMART; SM00282; LamG; 1.
DR   PROSITE; PS00232; CADHERIN_1; 14.
DR   PROSITE; PS50268; CADHERIN_2; 32.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 1.
KW   Transmembrane; Glycoprotein; Repeat; EGF-like domain; Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19   4349       PROTOCADHERIN FAT 2.
FT   DOMAIN      19   4048       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  4049   4069       POTENTIAL.
FT   DOMAIN    4070   4349       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      34    148       CADHERIN 1.
FT   DOMAIN     149    256       CADHERIN 2.
FT   DOMAIN     363    458       CADHERIN 3.
FT   DOMAIN     459    564       CADHERIN 4.
FT   DOMAIN     565    669       CADHERIN 5.
FT   DOMAIN     716    820       CADHERIN 6.
FT   DOMAIN     821    925       CADHERIN 7.
FT   DOMAIN     926   1032       CADHERIN 8.
FT   DOMAIN    1033   1137       CADHERIN 9.
FT   DOMAIN    1138   1242       CADHERIN 10.
FT   DOMAIN    1243   1346       CADHERIN 11.
FT   DOMAIN    1350   1448       CADHERIN 12.
FT   DOMAIN    1449   1555       CADHERIN 13.
FT   DOMAIN    1556   1660       CADHERIN 14.
FT   DOMAIN    1661   1758       CADHERIN 15.
FT   DOMAIN    1759   1872       CADHERIN 16.
FT   DOMAIN    1969   2070       CADHERIN 17.
FT   DOMAIN    2071   2171       CADHERIN 18.
FT   DOMAIN    2172   2272       CADHERIN 19.
FT   DOMAIN    2273   2379       CADHERIN 20.
FT   DOMAIN    2380   2481       CADHERIN 21.
FT   DOMAIN    2482   2585       CADHERIN 22.
FT   DOMAIN    2586   2691       CADHERIN 23.
FT   DOMAIN    2692   2797       CADHERIN 24.
FT   DOMAIN    2798   2906       CADHERIN 25.
FT   DOMAIN    2907   3011       CADHERIN 26.
FT   DOMAIN    3012   3113       CADHERIN 27.
FT   DOMAIN    3114   3218       CADHERIN 28.
FT   DOMAIN    3219   3321       CADHERIN 29.
FT   DOMAIN    3322   3426       CADHERIN 30.
FT   DOMAIN    3427   3531       CADHERIN 31.
FT   DOMAIN    3532   3642       CADHERIN 32.
FT   DOMAIN    3773   3944       LAMININ G-LIKE.
FT   DOMAIN    3951   3984       EGF-LIKE 1.
FT   DOMAIN    3989   4022       EGF-LIKE 2.
FT   DOMAIN    3297   3302       POLY-SER.
FT   CARBOHYD    39     39       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   210    210       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   280    280       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   568    568       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   627    627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   655    655       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   789    789       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   996    996       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1175   1175       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1303   1303       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1383   1383       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1417   1417       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1904   1904       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1998   1998       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2007   2007       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2165   2165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2183   2183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2325   2325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2368   2368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2387   2387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2430   2430       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2470   2470       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2547   2547       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2597   2597       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  2654   2654       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3125   3125       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3276   3276       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3310   3310       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3430   3430       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3471   3471       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3601   3601       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3772   3772       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3813   3813       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3840   3840       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3873   3873       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3904   3904       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  3989   3989       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT  4117   4117       P -> L (IN REF. 2).
FT   CONFLICT  4160   4160       E -> G (IN REF. 2).
SQ   SEQUENCE  4349 AA;  479383 MW;  6C050CA17AD398ED CRC64;

  Query Match             5.7%;  Score 92;  DB 1;  Length 4349;
  Best Local Similarity   21.9%;  Pred. No. 1.1e+02;
  Matches   80;  Conservative   41;  Mismatches  140;  Indels  104;  Gaps   16;

Qy        17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCK----------------------V 54
```

```
Db       2666 GGPPHWNSLVPVRLQVVPKKVSLPKF--SEPLYTFSAPEDLPEGSEIGIVKAVAAQDPVI 2723

Qy         55 DKLVSFIIPLAGSPGRFVVSLEREIAIL----TWDGVSAAPTSIEAIVN----------V 100

Db       2724 YSLVRGTTPESNKDG--VFSLDPDTGVIKVRKPMDHESTKLYQIDVMAHCLQNTDVVSLV 2781

Qy        101 EPHIKNNRLNDGK----ADPLGNLWTGTMAIDAGLPVGPVT---------GSL-YHLGAD 146

Db       2782 SVNIQVGDVNDNRPVFEADPYKAVLTENMPV--GTSVIQVTAIDKDTGRDGQVSYRLSAD 2839

Qy        147 KKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHE 206

Db       2840 PGSNVHEL-FAIDSESGWITTLQEL--------------DCETCQ--------TYHFH- 2874

Qy        207 VPGYPDGQTI------------DEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDP- 252

Db       2875 VVAYDHGQTIQLSSQALVQVSITDENDNAPRFASEEYRGSVVENSEPGELVATLKTLDAD 2934

Qy        253 ------QVTSVAFGGPNLDELHVTSAGLQLDDSS---LDKSLVNGHVYRVTGLGVKGFAG 303

Db       2935 ISEQNRQVTCYITEGDPLGQFGISQVGDEWRISSRKTLDREHTAKYLLRVTASDGKFQAS 2994

Qy        304 VKVKL 308

Db       2995 VTVEI 2999


RESULT 15
RL1_STRVG
ID   RL1_STRVG      STANDARD;      PRT;   242 AA.
AC   P48951;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   50S ribosomal protein L1.
GN   RPLA.
OS   Streptomyces virginiae.
OC   Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1961;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96257210; PubMed=8675024;
RA   Katayama M., Sakai Y., Okamoto S., Ihara F., Nihira T., Yamada Y.;
RT   "Gene organization in the ada-rplL region of Streptomyces virginiae.";
RL   Gene 171:135-136(1996).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC       LOCATED IN THE NEIGHBORHOOD OF THE SITE WHERE ELONGATION FACTOR TU
CC       IS BOUND TO THE RIBOSOME (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L1P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D50624; BAA09303.1; -.
DR   HSSP; P27150; 1AD2.
DR   InterPro; IPR002143; Ribosomal_L1.
DR   Pfam; PF00687; Ribosomal_L1; 1.
DR   ProDom; PD001314; Ribosomal_L1; 1.
DR   TIGRFAMs; TIGR01169; rplA_bact; 1.
DR   PROSITE; PS01199; RIBOSOMAL_L1; 1.
KW   Ribosomal protein; rRNA-binding.
SQ   SEQUENCE   242 AA;  26015 MW;  2194A6FF6FFB13CE CRC64;

  Query Match             5.7%;  Score 91.5;  DB 1;  Length 242;
  Best Local Similarity   26.8%;  Pred. No. 2.8;
  Matches   41;  Conservative   14;  Mismatches   63;  Indels   35;  Gaps    5;

Qy         25 ETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSLEREIAILTW 84

Db         31 ETSTTKFDGTVEVAFRLGVDPR------KADQMVRGTVNLPHGTGK----TARVLVFATG 80

Qy         85 DGVSAAPTSIEAIVNVEPHI----KNNRLND----------GKADPLG----------N 119

Db         81 DRAAAAEAAGADIVGDDELINEIAKGNRLNEFDAVVATPDLMGKVGRLGRVLGPRGLMPN 140

Qy        120 LWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMH 152

Db        141 PKTGTVTMDVAKAVTEIKGGKIEFRVDKHSNLH 173


Search completed: August 1, 2003, 14:32:45
Job time : 27 secs
```

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:26:38 ; Search time 70 Seconds
(without alignments)
906.607 Million cell updates/sec

Title: US-10-089-986-2
Perfect score: 1615
Sequence: 1 MGPVVEKIAELGKYTVGEGP.....HVYRVTGLGVKGFAGVKVKL 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1615	100.0	308	5	Q95YI4	Q95yi4 photinus py
2	889.5	55.1	309	5	Q8TA67	Q8ta67 luciola cru
3	859.5	53.2	307	5	Q8TA68	Q8ta68 luciola lat
4	568	35.2	303	5	Q9NDS6	Q9nds6 drosophila
5	568	35.2	303	5	Q9VFG5	Q9vfg5 drosophila
6	550	34.1	303	5	Q9VYR1	Q9vyr1 drosophila
7	538	33.3	306	5	Q9NDP1	Q9ndp1 sarcophaga
8	521	32.3	306	5	Q9BID6	Q9bid6 calliphora
9	474.5	29.4	299	13	Q9I922	Q9i922 xenopus lae
10	470.5	29.1	299	13	Q9IBA8	Q9iba8 xenopus lae
11	455	28.2	299	13	Q9I923	Q9i923 gallus gall
12	453.5	28.1	299	11	Q925W3	Q925w3 rattus norv
13	445.5	27.6	267	5	Q95PD9	Q95pd9 calliphora
14	423.5	26.2	569	16	Q92ZR8	Q92zr8 rhizobium m
15	409.5	25.4	584	16	Q8UJL2	Q8ujl2 agrobacteri
16	397.5	24.6	271	17	Q96ZI9	Q96zi9 sulfolobus
17	365.5	22.6	285	17	Q97VC7	Q97vc7 sulfolobus
18	350.5	21.7	282	17	Q9HKF3	Q9hkf3 thermoplasm
19	350	21.7	285	17	Q96XG1	Q96xg1 sulfolobus
20	348	21.5	211	5	Q8T0Z5	Q8t0z5 sarcophaga
21	339.5	21.0	286	16	Q9L0A7	Q910a7 streptomyce
22	331	20.5	289	16	Q9A9Z1	Q9a9z1 caulobacter
23	315	19.5	293	16	Q9A7B9	Q9a7b9 caulobacter
24	308	19.1	282	17	Q97A53	Q97a53 thermoplasm
25	294.5	18.2	303	16	Q8XVS4	Q8xvs4 ralstonia s
26	292	18.1	285	17	Q97UH7	Q97uh7 sulfolobus
27	288.5	17.9	311	16	Q931B6	Q931b6 rhizobium m
28	287.5	17.8	296	16	Q8XRJ9	Q8xrj9 ralstonia s
29	275.5	17.1	294	16	Q92T70	Q92t70 rhizobium m
30	273	16.9	322	16	Q8U8A8	Q8u8a8 agrobacteri
31	270	16.7	301	16	Q986F6	Q986f6 rhizobium l
32	256	15.9	314	16	Q8UHI0	Q8uhi0 agrobacteri
33	228.5	14.1	294	16	Q92RN9	Q92rn9 rhizobium m
34	213	13.2	302	16	Q8ZMW6	Q8zmw6 salmonella
35	206	12.8	305	16	Q8U8X6	Q8u8x6 agrobacteri
36	203	12.6	302	16	Q8Z4I8	Q8z4i8 salmonella
37	202.5	12.5	291	16	Q8YE81	Q8ye81 brucella me
38	201	12.4	290	16	Q98DE0	Q98de0 rhizobium l
39	153.5	9.5	348	16	Q8U8R1	Q8u8r1 agrobacteri
40	145.5	9.0	296	16	Q984W1	Q984w1 rhizobium l
41	140	8.7	303	16	Q926B8	Q926b8 rhizobium m
42	132	8.2	304	16	Q926I2	Q926i2 rhizobium m
43	126.5	7.8	115	11	Q9QWP2	Q9qwp2 rattus norv
44	124	7.7	353	16	Q8XRK7	Q8xrk7 ralstonia s
45	122.5	7.6	306	16	Q98EA9	Q98ea9 rhizobium l

ALIGNMENTS

RESULT 1
Q95YI4
ID Q95YI4 PRELIMINARY; PRT; 308 AA.
AC Q95YI4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Luciferin regenerating enzyme.
OS Photinus pyralis (North American firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Photinus.
OX NCBI_TaxID=7054;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21453302; PubMed=11457857;
RA Gomi K., Kajiyama N.;
RT "Oxyluciferin, a Luminescence Product of Firefly Luciferase, Is
RT Enzymatically Regenerated into Luciferin.";
RL J. Biol. Chem. 276:36508-36513(2001).
DR EMBL; AB062786; BAB60700.1; -.
SQ SEQUENCE 308 AA; 33806 MW; 4403A6EC6459A378 CRC64;

Query Match 100.0%; Score 1615; DB 5; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.8e-125;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNL 120

Qy        121 WTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 WTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRR 180
```

```
Qy        181 VDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQP 240

Qy        241 EVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG 300

Qy        301 FAGVKVKL 308
              ||||||||
Db        301 FAGVKVKL 308


RESULT 2
Q8TA67
ID   Q8TA67      PRELIMINARY;      PRT;   309 AA.
AC   Q8TA67;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Luciferin-regenerating enzyme.
GN   G-LRE.
OS   Luciola cruciata (Japanese firefly) (Genji firefly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC   Elateriformia; Cantharoidea; Lampyridae; Luciola.
OX   NCBI_TaxID=7051;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gomi K., Hirokawa K., Kajiyama N.;
RT   "Molecular cloning and expression of luciferin-regenerating enzyme
RT   (LRE) cDNA in Luciola cruciata and Luciola lateralis.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB072448; BAB85479.1; -.
SQ   SEQUENCE   309 AA;  33804 MW;  BE33C4E21EF51E6B CRC64;

  Query Match           55.1%;  Score 889.5;  DB 5;  Length 309;
  Best Local Similarity  56.6%;  Pred. No. 1.8e-65;
  Matches  176;  Conservative   47;  Mismatches   83;  Indels    5;  Gaps    4;

Qy          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60
              | | ||:| ||| | : | |||| |||:||||| | :: :|||: | :|  | ||  ||
Db          1 MAPTVEQIVELGTYLLAESPHWDDETQSLYFVDIVGRSVNKYVPTTKTHTQLKFDKNPSF 60

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIE--AIVNVEPHIKNNRLNDGKADPLG 118
              |||: |   ||:|||||| :|||| |:||: ||  |: :  |    ||||||||||||
Db         61 IIPVKGCSDRFIVSLEREINLLTWDGASSAPSKIEKIAVFDNTPEKSENRLNDGKADPLG 120

Qy        119 NLWTGTMAIDAGLPVG-PVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSG 177
              ||| ||| : :    | || |:|  | ::|:|| | | : |:|||||| |||| |||||
Db        121 NLWAGTMNMGSDHTTGTPVRGTLSSL-SNKQVKEHVSEVCISNGLAWSKDLKKFYYIDSA 179

Qy        178 KRRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKIST 237
               |:||::|:||  ||:|||::|||||:|| : | ||||||| |||||||  || :::|| |
Db        180 VRQVDQFDFDAKNLSLSNRQPLFTFDKHGIMGSPDGQTIDAEGNLWVATCQGDKVLKIDT 239

Qy        238 QQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLG 297
              || ||  |:||: ||||| || |: |:||:| ::|  :   | :  | :|:|||||
Db        240 STPETLLGIVEIPEHQVTSVCIGGAELNVLYVTTASIKLPGADETKPM-KGAIYKVTGLG 298

Qy        298 VKGFAGVKVKL 308
              |||  | :|||
Db        299 VKGLPGDRVKL 309


RESULT 3
Q8TA68
ID   Q8TA68      PRELIMINARY;      PRT;   307 AA.
AC   Q8TA68;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Luciferin-regenerating enzyme.
GN   H-LRE.
OS   Luciola lateralis (Firefly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC   Elateriformia; Cantharoidea; Lampyridae; Luciola.
OX   NCBI_TaxID=7052;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gomi K., Hirokawa K., Kajiyama N.;
RT   "Molecular cloning and expression of luciferin-regenerating enzyme
RT   (LRE) cDNA in Luciola cruciata and Luciola lateralis.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB072447; BAB85478.1; -.
SQ   SEQUENCE   307 AA;  34286 MW;  B03CED0AC609A92C CRC64;

  Query Match           53.2%;  Score 859.5;  DB 5;  Length 307;
  Best Local Similarity  52.8%;  Pred. No. 5.5e-63;
  Matches  163;  Conservative   58;  Mismatches   85;  Indels    3;  Gaps    3;

Qy          1 MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSF 60
              | ||:|:| |:  : :||||||| |||:||||| :||: |||||| |::|   ::|  ||
Db          1 MSPVIEQITEVDNFQIGEGPHWDTETQSLYFVDILEKSIHKYVPSTKQHTKMILNKRPSF 60

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNL 120
              |||:  :  |||:||||:| :|||||||| |: :| || |:  |: | ||||||  |||
Db         61 IIPIKETSDRFVISLERDICVLTWDGVSATPSHLETIVTVDTGIEGNTFNDGKADAFGNL 120

Qy        121 WTGTMAIDAGL-PVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKR 179
              | ||:     :   || ||:|| | ::|::: | ||| ::|||||: | || |:||| ||
Db        121 WAGTLYSKFDIEKQGPNTGTLYSL-SNKQLRKHISNIFLSNGLAWNKDSKKFYFIDSNKR 179

Qy        180 RVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQ 239
               :|::|||:  | ||| :|||| :|| : | || |||||  |||||: :| ::| | |:|
Db        180 TIDQFDYDSENLIISNCQPLFTLDKHGIQGLPDAQTIDENDNLWVAIVRGGKVINIGTKQ 239

Qy        240 PEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVK 299
              || ||  : :|:  :||| ||| ||||:||::|:  ::   | || | :|||||||||
Db        240 PESLLGVINMPESLITSVCFGGSKLDELYVTTSGIKEYETDSTK-LVKGGLYRVTGLGVK 298

Qy        300 GFAGVKVKL 308
              |    :  |
Db        299 GLPAHRFSL 307


RESULT 4
Q9NDS6
ID   Q9NDS6      PRELIMINARY;      PRT;   303 AA.
AC   Q9NDS6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   Senescence marker protein-30 (SMP-30).
GN   SMP-30 OR CG7390.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON S; TISSUE=WHOLE BODY;
RX   PubMed=10817837;
RA   Goto S.G.;
RT   "Expression of Drosophila homologue of senescence marker protein-30
RT   during cold acclimation.";
RL   J. Insect Physiol. 46:1111-1120(2000).
DR   EMBL; AB029490; BAA92938.1; -.
DR   FlyBase; FBgn0038257; smp-30.
SQ   SEQUENCE   303 AA;  33349 MW;  68F12B92E1AAB5AD CRC64;

  Query Match           35.2%;  Score 568;  DB 5;  Length 303;
  Best Local Similarity  40.7%;  Pred. No. 6.9e-39;
  Matches  121;  Conservative   48;  Mismatches  116;  Indels   12;  Gaps    5;
```

```
Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDK--LVSFIIPLAGSPGRFVV 73
Db         15 LGEGPHWDVDRQSLYYVDLESAGINRYDFKQNKVYKAKIEGEIFASFILPVENKPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
Db         75 GCGLRTVIVQWDGVSAVAKVTRTLFEVQPDLKENRLNDAKTDPNGRFYGGTMA-DSGDIF 133

Qy        134 GPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSI 193
Db        134 TQWKGELYSWQAGGQPNAIRSKVGISNGLAWDVKAKKFYFIDTNNHEVLAYDYNQSTGAV 193

Qy        194 SNQRPLFTFEK--HEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPD 251
Db        194 SNPKVIFDLRKIRPEGPLFPDGMTVDTDGNIYVATFNGGTVFKVNPSTGKILLE-IKIPT 252

Qy        252 PQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVKL 308
Db        253 TQITSVAFGGPNLDILYVTTA------NKFDQPKPAGTTFQVTGLNAKGYAGVNLKI 303
```

RESULT 5
Q9VFG5
ID Q9VFG5 PRELIMINARY; PRT; 303 AA.
AC Q9VFG5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG7390 protein (RE17477p).
GN SMP-30 OR CG7390.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE003706; AAF55095.1; -.
DR EMBL; AY071130; AAL48752.1; -.
DR FlyBase; FBgn0038257; smp-30.
SQ SEQUENCE 303 AA; 33377 MW; 68F92692E1AC88AD CRC64;

Query Match 35.2%; Score 568; DB 5; Length 303;
Best Local Similarity 40.7%; Pred. No. 6.9e-39;
Matches 121; Conservative 48; Mismatches 116; Indels 12; Gaps 5;

```
Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDK--LVSFIIPLAGSPGRFVV 73
Db         15 LGEGPHWDVDRQSLYYVDLESAGINRYDFKQNKVYRAKIEGEIFASFILPVENKPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
Db         75 GCGLRTVIVQWDGVSAVAKVTRTLFEVQPDLKENRLNDAKTDPNGRFYGGTMA-DSGDIF 133

Qy        134 GPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSI 193
Db        134 TQWKGELYSWQAGGQPNAIRSKVGISNGLAWDVKAKKFYFIDTNNHEVLAYDYNQSTGAV 193

Qy        194 SNQRPLFTFEK--HEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPD 251
Db        194 SNPKVIFDLRKIRPEGPLFPDGMTVDTDGNIYVATFNGGTVFKVNPSTGKILLE-IKIPT 252

Qy        252 PQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVKL 308
Db        253 TQITSVAFGGPNLDILYVTTA------NKFDQPKPAGTTFQVTGLNAKGYAGVNLKI 303
```

RESULT 6
Q9VYR1
ID Q9VYR1 PRELIMINARY; PRT; 303 AA.
AC Q9VYR1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CG1803 protein (Regucalcin HOMOLOGUE).
GN REGUCALCIN OR CG1803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

```
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20250851; PubMed=10788801;
RA   Nakajima Y., Natori S.;
RT   "Identification and characterization of an anterior fat body protein
RT   in an insect.";
RL   J. Biochem. 127:901-908(2000).
DR   EMBL; AE003487; AAF48129.1; -.
DR   EMBL; AB036904; BAA99283.1; -.
DR   FlyBase; FBgn0030362; regucalcin.
SQ   SEQUENCE   303 AA;  33644 MW;  F9AD6EB15E758209 CRC64;

  Query Match             34.1%;  Score 550;  DB 5;  Length 303;
  Best Local Similarity   39.2%;  Pred. No. 2.1e-37;
  Matches  118;  Conservative   48;  Mismatches  115;  Indels   20;  Gaps    6;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSFIIPLAGSPGRFVV 73
              :|||||||   |:||:||    :  :|  :| |    |:  : |  |::|: | |  | |
Db         15 LGEGPHWDVARQSLYYVDLEAGSLLRYDYAQNKVYKTKIEGETLAGFVLPVEGRPQEFAV 74

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
                | : |: ||||| :   :  :  |:| :: ||||| | || |  : |||      :
Db         75 GCGRRVVIVNWDGVSPSAKVVRTLFEVQPLMEKNRLNDAKVDPRGRFFGGTMRY-----I 129

Qy        134 GPV----TGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDAS 189
                |        | ||    |  :| : : :: |:|||||    || ||||:    |  ||||
Db        130 GDEFEFRHGELYRWEAGGQVSVIKGDVGISNGLAWDEKAKKFYYIDTTDYEVKSYDYDFE 189

Qy        190 TLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTV 247
                |   || : :|   |:    :  ||| ||| ||||:|| | |  | |::    ::||: :
Db        190 TGVASNPKVIFNLRKNSPKDHLLPDGLTIDTEGNLYVATFNGATIYKVNPNTGKILLE-I 248

Qy        248 KIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFAGVKVK 307
                | |  |:|| |||||||| |:||:|      :  |:    |  |:||||   |: || :|
Db        249 KFPTKQITSAAFGGPNLDILYVTTA------AKFDQPAPAGTTYKVTGLNATGYPGVNLK 302

Qy        308 L 308
                :
Db        303 V 303


RESULT 7
Q9NDP1
ID   Q9NDP1      PRELIMINARY;      PRT;   306 AA.
AC   Q9NDP1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   Anterior fat body protein.
GN   AFP.
OS   Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Sarcophagidae; Sarcophaga.
OX   NCBI_TaxID=7386;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=FAT BODY;
RX   MEDLINE=20250851; PubMed=10788801;
RA   Nakajima Y., Natori S.;
RT   "Identification and characterization of an anterior fat body protein
RT   in an insect.";
RL   J. Biochem. 127:901-908(2000).
DR   EMBL; AB036903; BAA99282.1; -.
SQ   SEQUENCE   306 AA;  34319 MW;  77CFF54575D7E8E7 CRC64;

  Query Match             33.3%;  Score 538;  DB 5;  Length 306;
  Best Local Similarity   39.0%;  Pred. No. 2.1e-36;
  Matches  122;  Conservative   48;  Mismatches  123;  Indels   20;  Gaps    7;

Qy          3 PVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSF 60
              |: :  |||     |||||||  || ||:||       :|  :: |    |:  : |  |
Db          7 PLPDSHAEL-----GEGPHWDVATQNLYYVDINAGKLLRYNYNENKVYKAKIEGEDLAGF 61

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNV---EPHIKNNRLNDGKADPL 117
               ||: |:  :| |    | : |: |||:|     |: :  |   :      || |||| ||
Db         62 AIPVEGTTDQFAVGSGRRVVIVQWDGISETAKVIKTLFEVQTGDERFNGNRFNDGKCDPR 121

Qy        118 GNLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSG 177
              | |: |||    |       | ||      |::: :||: |:|||||:   || ||||:
Db        122 GRLFAGTMRY-VGDEFEHRYGELYKYEKGGKIEVVKSNVGISNGLAWNEKTKKFYYIDTT 180

Qy        178 KRRVDEYDYDASTLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKI 235
                 | |||||  |   :| : :|   |:    :  ||| ||| :||::|| | |  | |:
Db        181 DYEVKEYDYDFETGKATNPKVVFNLRKNSPKDHLLPDGMTIDTDGNIYVATFNGYTIYKV 240

Qy        236 STQQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTG 295
              :   :|||: :| |  |:|| |||||||| |:||::      |   :    |  |:|||
Db        241 NPTTGKVLLE-IKFPCKQITSAAFGGPNLDILYVTTS------SRFGEPDPAGTTYKVTG 293

Qy        296 LGVKGFAGVKVKL 308
              || ||:   |::|
Db        294 LGAKGYPMTKIQL 306


RESULT 8
Q9BID6
ID   Q9BID6      PRELIMINARY;      PRT;   306 AA.
AC   Q9BID6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Anterior fat body protein.
OS   Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=7373;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ANTERIOR FAT BODY;
RA   Hansen I.A., Meyer S.R., Schaefer I., Scheller K.;
RT   "Identification of the anterior fat body protein as interactor of the
RT   hexamerin receptor in the blowfly Calliphora vicina.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY028616; AAK26174.1; -.
SQ   SEQUENCE   306 AA;  34289 MW;  497241A2424C2C19 CRC64;
```

```
  Query Match          32.3%;  Score 521;  DB 5;  Length 306;
  Best Local Similarity   39.0%;  Pred. No. 5.4e-35;
  Matches  122;  Conservative   44;  Mismatches  127;  Indels   20;  Gaps   7;

Qy          3 PVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKV--DKLVSF 60
              |: :  |||     |||||||  || ||:||       :|  :: |    |:  : |  |
Db          7 PLPDSHAEL-----GEGPHWDVATQNLYYVDINAGKLLRYNYNENKVYKAKIEGEDLAGF 61

Qy         61 IIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNV---EPHIKNNRLNDGKADPL 117
              :|: |:  :| |   | : |: |||||     :: :  |   :      || |||| ||
Db         62 AVPVEGTTDQFAVGSGRRVVIVQWDGVSETAKVLKTLFEVQTGDERFTGNRFNDGKCDPR 121

Qy        118 GNLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSG 177
              | |: |||    |       | ||      :|:: :|:: |:|||||:   || ||||:
Db        122 GRLFAGTMRY-VGDEFEHRYGELYKYENGGEVEVIKSDVGISNGLAWNEKTKKFYYIDTT 180

Qy        178 KRRVDEYDYDASTLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQGQRIIKI 235
                | |||||  |   || : :|   |     :  ||| ||| ||||:|| | |  | ||
Db        181 DYEVKEYDYDFETGKASNPKVVFNLRKTSPKDHLLPDGMTIDTEGNLYVATFNGHSIYKI 240

Qy        236 STQQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTG 295
              :   :|||: :| |  |:|| ||||||   | ||::     |   :    |  |:|||
Db        241 NPTTGQVLLE-IKFPCKQITSAAFGGPNFHILFVTTS------SRFGEPHPAGTTYKVTG 293

Qy        296 LGVKGFAGVKVKL 308
              || ||:   |::|
Db        294 LGAKGYPMTKIQL 306


RESULT 9
Q9I922
ID   Q9I922      PRELIMINARY;      PRT;   299 AA.
AC   Q9I922;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   Regucalcin.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20351777; PubMed=10891565;
RA   Misawa H., Yamaguchi M.;
RT   "The gene of Ca2+-binding protein regucalcin is highly conserved in
RT   vertebrate species.";
RL   Int. J. Mol. Med. 6:191-196(2000).
DR   EMBL; AB037936; BAA90694.1; -.
SQ   SEQUENCE   299 AA;  33087 MW;  BABB9749258BB2C7 CRC64;

  Query Match          29.4%;  Score 474.5;  DB 13;  Length 299;
  Best Local Similarity   37.2%;  Pred. No. 3.7e-31;
  Matches  110;  Conservative   46;  Mismatches  117;  Indels   23;  Gaps   7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              | :|| | |: :  || |||    :   :: || ||     |:  :  :        | :|:
Db         14 YKIGESPVWEEKEGTLLFVDITGQKVCRWDPSTKKVQSVSVEAPIGSV--ALRKSGGYVL 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::    : | |:       |:  :  |:    ||| |||| || |    |||: :    :
Db         72 AMGNTFSALNWE-----DQSVTTLARVDEDKPNNRFNDGKVDPEGRFLAGTMSQE----I 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
              |       |||: |  |  |  |   : |:||| || | | :|||||   :||  |||
Db        123 RPAVVERNQGSLFTLYPDHSVVKHFDMVDISNGLDWSLDHKTLYYIDSLSFKVDALDYDM 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
              |   ||:| |:  :: |  | |||  || || |||| : | |:|:|    ||    : |
Db        183 KTGKSSNRRTLYKLQQDE--GIPDGMCIDAEGKLWVACYNGGRVIRI---DPETGKQIQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||||:  |::||||   :|:    :   :| :|:::|||||| |
Db        238 VKLPIDKTTSCCFGGPDYSEMYVTSACDGMDEDWKKRQPQSGGIYKITGLGVKGIA 293


RESULT 10
Q9IBA8
ID   Q9IBA8      PRELIMINARY;      PRT;   299 AA.
AC   Q9IBA8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   Senescence marker protein-30.
GN   XSMP-30.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ANIMAL CAP;
RX   MEDLINE=20193502; PubMed=10727865;
RA   Sato A., Asashima M., Yokota T., Nishinakamura R.;
RT   "Cloning and expression pattern of a Xenopus pronephros-specific gene,
RT   XSMP-30.";
RL   Mech. Dev. 92:273-275(2000).
DR   EMBL; AB033368; BAA93719.1; -.
SQ   SEQUENCE   299 AA;  33073 MW;  BABB9749258BB41C CRC64;

  Query Match          29.1%;  Score 470.5;  DB 13;  Length 299;
  Best Local Similarity   36.8%;  Pred. No. 7.9e-31;
  Matches  109;  Conservative   47;  Mismatches  117;  Indels   23;  Gaps   7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              | :|| | |: :  || |||    :   :: || ||     |:  :  :        | :|:
Db         14 YKIGESPVWEEKEGTLLFVDITGQKVCRWDPSTKKVQSVSVEAPIGSV--ALRKSGGYVL 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::    : | |:       |:  :  |:    ||| |||| || |    |||: :    :
Db         72 AMGNTFSALNWE-----DQSVTTLARVDEDKPNNRFNDGKVDPEGRFLAGTMSQE----I 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
              |       |||: |  |  |  |   : |:||| || | | :|||||   :||  |||
Db        123 RPAVVERNQGSLFTLYPDHSVVKHFDMVDISNGLDWSLDHKTLYYIDSLSFKVDALDYDM 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
              |   ||:| |:  :: |  | |||  || || |||| : | |:|:|    ||    : |
Db        183 KTGKSSNRRTLYKLQQDE--GIPDGMCIDAEGKLWVACYNGGRVIRI---DPETGKQIQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||||:  |::||||   :|:    :   :| :|:::|||||| |
Db        238 VKLPIDKTTSCCFGGPDYSEMYVTSACDGMDEDWKKRQPQSGGIYKISGLGVKGIA 293


RESULT 11
Q9I923
ID   Q9I923      PRELIMINARY;      PRT;   299 AA.
AC   Q9I923;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   Regucalcin.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=20351777; PubMed=10891565;
RA   Misawa H., Yamaguchi M.;
```

```
RT   "The gene of Ca2+-binding protein regucalcin is highly conserved in
RT   vertebrate species.";
RL   Int. J. Mol. Med. 6:191-196(2000).
DR   EMBL; AB037935; BAA90693.1; -.
SQ   SEQUENCE   299 AA;  33229 MW;  4754C7571164720E CRC64;

  Query Match             28.2%;  Score 455;  DB 13;  Length 299;
  Best Local Similarity   37.0%;  Pred. No. 1.5e-29;
  Matches  112;  Conservative   40;  Mismatches  123;  Indels   28;  Gaps    8;

Qy         13 KYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFV 72
              :| :|| | || :  :|  ||   :   ::  :  :     ||  || :      | :|
Db         13 RYRLGESPVWDEKENSLLCVDITGRKVCRWDAASGQVQALSVDAPVSSV--ALRKSGDYV 70

Qy         73 VSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLP 132
              ::|    | | |         :  |  |:    ||| |||| || |. : |||| :
Db         71 ITLGTRFAALKW-----KEQLVTTIAQVDRDKANNRFNDGKVDPAGRYFAGTMAEE---- 121

Qy        133 VGPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYD 187
              : |       |||| |  |  |  |   : |:||| || | |  :|||    || :|||
Db        122 IRPAVLERRQGSLYTLCPDHSVVKHFDQVDISNGLDWSLDHKTFFYIDSLSYSVDAFDYD 181

Qy        188 ASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLD 245
              |  | |:| ::  || |    |||  || || |||| : | |:|::    ||    :
Db        182 LQTGKIGNRRSVYKLEKEE--SIPDGMCIDTEGKLWVACYDGGRVIRL---DPETGKRIQ 236

Qy        246 TVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKG----- 300
              |||:|  : ||  ||| :  |::||||   :|   | :    | |:::|||||||
Db        237 TVKLPVDKTTSCCFGGKDYSEMYVTSASDGMDREWLSRQPQAGGVFKITGLGVKGIPPYP 296

Qy        301 FAG 303
              |||
Db        297 FAG 299


RESULT 12
Q925W3
ID   Q925W3      PRELIMINARY;      PRT;   299 AA.
AC   Q925W3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Regucalcin.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Misawa H., Yamaguchi M.;
RT   "The gene family encoding the calcium-binding protein regucalcin.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037934; BAA90692.1; -.
SQ   SEQUENCE   299 AA;  33390 MW;  5D8F2D95FCA4EE35 CRC64;

  Query Match             28.1%;  Score 453.5;  DB 11;  Length 299;
  Best Local Similarity   37.5%;  Pred. No. 2e-29;
  Matches  111;  Conservative   40;  Mismatches  122;  Indels   23;  Gaps    7;

Qy         14 YTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV 73
              |  || | |:  :: | |||   ||  ::     :    ||  || :        | :|
Db         14 YRCGESPVWEEASKCLLFVDIPSKTVCRWDSISNRVQRVGVDAPVSSV--ALRQSGGYVA 71

Qy         74 SLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPV 133
              ::  :   | |:       |:  :  |:   |||| |||| || |  : |||| :
Db         72 TIGTKFCALNWE-----DQSVFILAMVDEDKKNNRFNDGKVDPAGRYFAGTMAEE----T 122

Qy        134 GPVT-----GSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDA 188
               |       |||| |  |  || :   : |:||| || | |  |||||    || :|||
Db        123 APAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSLDHKIFYYIDSLSYTVDAFDYDL 182

Qy        189 STLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEV--LLDT 246
               |  |||:| ::  || |    |||  || || |||| : | |:|::    ||    | |
Db        183 PTGQISNRRTVYKMEKDE--QIPDGMCIDVEGKLWVACYNGGRVIRL---DPETGKRLQT 237

Qy        247 VKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYRVTGLGVKGFA 302
              ||:|  : ||  ||| :  |::|| |   :    | :    |:::::|||||| |
Db        238 VKLPVDKTTSCCFGGKDYSEMYVTCARDGMSAEGLLRQPDAGNIFKITGLGVKGIA 293


RESULT 13
Q95PD9
ID   Q95PD9      PRELIMINARY;      PRT;   267 AA.
AC   Q95PD9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Anterior fat body protein (Fragment).
GN   AFP.
OS   Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=7373;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=FAT BODY;
RA   Hansen I.A., Meyer S.R., Berlinger M.J., Scheller K.;
RT   "Identification of new interactors of the hexamerin receptor in the
RT   blowfly Calliphora vicina.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF326959; AAK51353.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   267 AA;  29836 MW;  76F4EEA957D032FF CRC64;

  Query Match             27.6%;  Score 445.5;  DB 5;  Length 267;
  Best Local Similarity   39.4%;  Pred. No. 7.8e-29;
  Matches  102;  Conservative   37;  Mismatches  107;  Indels   13;  Gaps    5;

Qy         55 DKLVSFIIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNV---EPHIKNNRLND 111
              : |  | :|: |:  :| |   | : |: |||||     :: :  |   : .   || ||
Db         17 EDLAGFAVPVEGTTDQFAVGSGRRVVIVQWDGVSETAKVLKTLFEVQTGDERFTGNRFND 76

Qy        112 GKADPLGNLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKM 171
              || || | |: |||     |        | ||      :|:: :|:: |:||||:   ||
Db         77 GKCDPRGRLFAGTMRY-VGDEFEHRYGELYKYENGGEVEVIKSDVGISNGLAWNEKTKKF 135

Qy        172 YYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVPGY--PDGQTIDEEGNLWVAVFQG 229
              ||||:    | |||||  |   || : :|   |      :  ||| ||| ||||:|| | |
Db        136 YYIDTTDYEVKEYDYDFETGKASNPKVVFNLRKTSPKDHLLPDGMTIDTEGNLYVATFNG 195

Qy        230 QRIIKISTQQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGH 289
              | ||:    :|||: :| |  |:|| ||||||   | ||::      |   :    |
Db        196 HSIYKINPTTGQVLLE-IKFPCKQITSAAFGGPNFHILFVTTS------SRFGEPHPAGT 248

Qy        290 VYRVTGLGVKGFAGVKVKL 308
               |:||||| ||:   |::|
Db        249 TYKVTGLGAKGYPMTKIQL 267


RESULT 14
Q92ZR8
ID   Q92ZR8      PRELIMINARY;      PRT;   569 AA.
AC   Q92ZR8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative calcium binding transcriptional regulatory protein.
GN   RA0381 OR SMA0717.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymA (megaplasmid 1).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=1021;
RX   MEDLINE=21396509; PubMed=11481432;
RA   Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA   Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA   Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA   Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA   Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT   "Nucleotide sequence and predicted functions of the entire
RT   Sinorhizobium meliloti pSymA megaplasmid.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR   EMBL; AE007229; AAK65039.1; -.
DR   InterPro; IPR000285; HTH_IclR.
DR   Pfam; PF01614; IclR; 1.
KW   Plasmid; Hypothetical protein; Complete proteome.
SQ   SEQUENCE   569 AA;  60943 MW;  EB34C90D654EE8AD CRC64;

  Query Match             26.2%;  Score 423.5;  DB 16;  Length 569;
  Best Local Similarity   34.6%;  Pred. No. 1.6e-26;
  Matches  100;  Conservative   46;  Mismatches  102;  Indels   41;  Gaps    7;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75
              :|| | |    : ||:|| :    |:: |   |   | | |||| ::|          :
Db        293 LGEAPVWVEREKRLYWVDILHPAVHRFDPVTGKNESCNVAKLVSAVLP----------TR 342

Qy         76 EREIAILTWDGVS------------AAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTG 123
                : : : |||             | |        || :  ||||| | || | || |
Db        343 NEGLIVASQDGVEHFDFDRGDFNPFAEP---------EPGLPENRLNDAKVDPSGRLWVG 393

Qy        124 TMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDE 183
              :| :|   |    ||||| | :  :|    |   :|||||| |    |::|:    :
Db        394 SMRLDVSRP----TGSLYRLTSAGEVTRAGSGFTVANGLAWSPDSSTFYFVDTVPGIIYA 449

Qy        184 YDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVL 243
              ||:||   ||:|:|   |    |  | ||| .:| :| :| |:: | |   ::  :|:
Db        450 YDFDAREGSIANRRVFVTVP--EAEGRPDGLAVDADGGVWCAIWDGWR---VNRYRPDGR 504

Qy        244 LD-TVKIPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVY 291
              ||  |::| |: |||||||  |  | :|||  :|  |:| :: ::| ::
Db        505 LDRAVELPVPRPTSVAFGGDELATLFITSARTRLPASTLTEAPLSGGIF 553


RESULT 15
Q8UJL2
ID   Q8UJL2      PRELIMINARY;      PRT;   584 AA.
AC   Q8UJL2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Transcriptional regulator, IclR family/regucalcin.
GN   ATU5464 OR AGR_PAT_685.
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OG   Plasmid AT.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE008967; AAL46151.1; ALT_INIT.
DR   EMBL; AE007915; AAK90839.1; -.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   584 AA;  62034 MW;  7E9F274048791037 CRC64;

  Query Match             25.4%;  Score 409.5;  DB 16;  Length 584;
  Best Local Similarity   33.1%;  Pred. No. 2.4e-25;
  Matches   99;  Conservative   57;  Mismatches  110;  Indels   33;  Gaps    9;

Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75
              :|| | |      ||:|| :    |:: |:  :   |:  |||| :||: |  || :|:
Db        308 LGESPVWHEGEDALYWVDILHPAVHRFDPATGRNETCETGKLVSAVIPVTG--GRLLVAS 365

Qy         76 EREIAILTWDGV---SAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLP 132
              :        |||      |   :   |: | | :|||| |  | | :| |:| |||  |
Db        366 Q--------DGVEWLDFASGRLTPFVSPEAGIADNRLNDAKCGPDGAIWVGSMRIDASKP 417

Qy        133 VGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLS 192
                ||:|| : |:   :  |  | ::||| || | :  |::|:    :  || | :| :
Db        418 ----TGALYRINANGASERKEGGIIVSNGLGWSPDGRTFYFVDTVPGLIHAYDCDPATGA 473

Qy        193 ISNQRPLFTFEKHEVP---GYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLD-TVK 248
              :| :|     |   :|   | |||  :| || :| |:: |  : :     |   ||  :
Db        474 LSQRR-----EFARIPVADGRPDGLAVDAEGGVWCAIWDGWCVRR---YLPNGKLDQVID 525

Qy        249 IPDPQVTSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGHVYR----VTGLGVKGFAG 303
              :| |: :|:|||||:|  | :|||  :|  |:|  : ::| ::     ::|  :  | |
Db        526 MPVPRPSSIAFGGPDLSTLFITSARTRLPASTLADAPLSGGLFSCRPGISGARIAMFEG 584


Search completed: August  1, 2003, 14:34:10
Job time : 76 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:32:13 ; Search time 24 Seconds
(without alignments)
377.594 Million cell updates/sec

Title: US-10-089-986-2
Perfect score: 1615
Sequence: 1 MGPVVEKIAELGKYTVGEGP.....HVYRVTGLGVKGFAGVKVKL 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	110	6.8	332	4	US-09-134-001C-4646	Sequence 4646, Ap
2	94	5.8	413	4	US-09-134-001C-5111	Sequence 5111, Ap
3	93	5.8	396	4	US-08-501-126-2	Sequence 2, Appli
4	91	5.6	397	2	US-08-282-197C-55	Sequence 55, Appl
5	90	5.6	1277	4	US-09-397-885-3	Sequence 3, Appli
6	89.5	5.5	387	1	US-08-112-630-2	Sequence 2, Appli
7	88.5	5.5	1529	4	US-09-134-001C-3945	Sequence 3945, Ap
8	88	5.4	1161	4	US-09-327-536-2	Sequence 2, Appli
9	86.5	5.4	428	4	US-09-548-372D-51	Sequence 51, Appl
10	86.5	5.4	428	4	US-09-548-367D-51	Sequence 51, Appl
11	86.5	5.4	434	4	US-09-548-372D-53	Sequence 53, Appl
12	86.5	5.4	434	4	US-09-548-367D-53	Sequence 53, Appl
13	86.5	5.4	476	4	US-09-548-372D-6	Sequence 6, Appli
14	86.5	5.4	476	4	US-09-548-372D-73	Sequence 73, Appl
15	86.5	5.4	476	4	US-09-548-367D-6	Sequence 6, Appli
16	86.5	5.4	476	4	US-09-548-367D-73	Sequence 73, Appl
17	86.5	5.4	526	4	US-09-342-648-8	Sequence 8, Appli
18	84	5.2	1112	2	US-08-714-402-2	Sequence 2, Appli
19	83.5	5.2	409	4	US-09-134-001C-3790	Sequence 3790, Ap
20	83	5.1	758	4	US-09-134-001C-4588	Sequence 4588, Ap
21	81.5	5.0	568	1	US-08-528-199-1	Sequence 1, Appli
22	81.5	5.0	1579	3	US-08-755-587-184	Sequence 184, App
23	81	5.0	287	3	US-09-173-581-2	Sequence 2, Appli
24	81	5.0	287	4	US-09-420-915-2	Sequence 2, Appli
25	81	5.0	288	4	US-09-441-039-2	Sequence 2, Appli
26	80.5	5.0	409	4	US-09-071-035-316	Sequence 316, App
27	80.5	5.0	433	4	US-09-071-035-314	Sequence 314, App
28	80.5	5.0	2329	3	US-08-755-587-16	Sequence 16, Appl
29	80	5.0	505	3	US-08-747-221B-14	Sequence 14, Appl
30	80	5.0	505	4	US-09-005-051-14	Sequence 14, Appl
31	80	5.0	550	3	US-08-747-221B-58	Sequence 58, Appl
32	80	5.0	550	4	US-09-005-051-58	Sequence 58, Appl
33	80	5.0	755	4	US-09-342-648-2	Sequence 2, Appli
34	80	5.0	932	4	US-08-887-534A-45	Sequence 45, Appl
35	79.5	4.9	540	4	US-09-134-001C-4705	Sequence 4705, Ap
36	79	4.9	305	3	US-08-965-600-1	Sequence 1, Appli
37	79	4.9	305	4	US-09-489-506-1	Sequence 1, Appli
38	79	4.9	474	2	US-08-836-854-9	Sequence 9, Appli
39	79	4.9	530	3	US-08-747-221B-53	Sequence 53, Appl
40	79	4.9	530	4	US-09-005-051-53	Sequence 53, Appl
41	79	4.9	550	3	US-08-747-221B-19	Sequence 19, Appl
42	79	4.9	550	4	US-09-005-051-19	Sequence 19, Appl
43	79	4.9	581	4	US-09-393-627B-1	Sequence 1, Appli
44	79	4.9	593	1	US-07-961-522-4	Sequence 4, Appli
45	79	4.9	593	1	US-08-217-438-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-134-001C-4646
; Sequence 4646, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4646
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4646

Query Match 6.8%; Score 110; DB 4; Length 332;
Best Local Similarity 24.5%; Pred. No. 0.00084;
Matches 58; Conservative 40; Mismatches 111; Indels 28; Gaps 11;

```
Qy          7 KIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAG 66
              ||:: |     || ::: | | |: :|  |    |  |: |: |     |  |    :
Db         44 KISDKGLQL--EGLNFNREGQ-LFLLDVFEGNIFKVNPATKEVT----TKFQS----VKD 92

Qy         67 SPGRFVVSLEREIAIL------TWDGVSAAPTSIEAIVNVEPHIKNNR-LNDGKADPLGN 119
              :|    |  :  : |       |  |: |     | |  :   :     ::|   | |
Db         93 NPAAIKVHKDGRLFICYLGDFKTTGGIFATTEKGEQIEEIISDLNTEYCIDDMVFDSKGG 152

Qy        120 LWTGTMAIDAGLPVGPVTGSLYHLGAD-KKVKMHESNIAIANGLAWSNDLKKMYYIDSGK 178
               :        |    |: | :|::  | | |     ||:::|||:| | | | ::  ::
Db        153 FYFTDF---RGYSTQPL-GGVYYVDPDFKTVTPIIQNISVANGIALSTDEKVLWVTETTT 208

Qy        179 RRVDE--YDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRII 233
                |:     : |  |::       : |  ||    ||   ||   ||:||:|:   |::
Db        209 NRLHRIALEDDGVTIAPFGATIPYYFTGHE---GPDSCCIDSNDNLYVAMYGQGRVL 262
```

RESULT 2
US-09-134-001C-5111
; Sequence 5111, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

```
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5111
;   LENGTH: 413
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5111

  Query Match             5.8%;  Score 94;  DB 4;  Length 413;
  Best Local Similarity   22.6%;  Pred. No. 0.074;
  Matches   76;  Conservative   46;  Mismatches  101;  Indels  114;  Gaps   17;

Qy         36 EKTFHKY-VPSQKKYTFCKVDKLVSFIIPLAGSPGRFVV----------SLEREIAIL-- 82
              ||  | |   |||::||  :  ::  ||  ||  |  :|          | | :
Db          6 EKNQHHYRNSSQKRHTFPWIKTIIVAII--AGIIGALLVLGIGKLLNKTGFNNEGATVHQ 63

Qy         83 --------TWDGVSAAPTSI-EAIVNVEPHIKNNRLNDGKADPLGNLWTG--TMAIDAGL 131
                     || |    |: : | :| | |    :|  |:  | :|: |  : : :||:
Db         64 VSNSHGGNQLDGKSNQYKSVHDMIKDVSPAIV-GVINMQKSTNLDDLFNGKASKSKEAGI 122

Qy        132 PVGPV----TGSLY-----HL--GADK-KVKMHESN---------------------- 155
               | :     || |     |:  || : ||::| |
Db        123 GSGVIYQISDGSAYIVTNNHVVDGASEIKVQLHNSKQVDAKLIGKDALTDIAVLKIKDTK 182

Qy        156 ----IAIAN----------------GLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSIS 194
                  |  ||                || ::|     : ||     |   ||:| : :
Db        183 GIKAIQFANSSKVQTGDSVFAMGNPLGLEFANS------VTSGIISASERTIDANTSAGN 236

Qy        195 NQRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDPQV 254
               :  :   :    ||   |  :|  |||              | ::::||   ||
Db        237 TKVNVLQTDAAINPGNSGGALVDINGNL----------------VGINSMKIAAAQV 277

Qy        255 TSVAFGGPNLDELHVTSAGL----QLDDSSLDKSLVN 287
               : |  |: :|: ||   |    :::  |:   |:|
Db        278 EGIGFAIPS-NEVRVTIEQLVKHGKIERPSIGIGLIN 313


RESULT 3
US-08-501-126-2
; Sequence 2, Application US/08501126
; Patent No. 6140095
;  GENERAL INFORMATION:
;    APPLICANT:  Van Solingen, Pieter
;    APPLICANT:  Williams, Diane P.
;    APPLICANT:  Iverson, Sara
;    APPLICANT:  Farrell, Roberta L.
;    APPLICANT:  Herbes, Wilhelmina T.
;    APPLICANT:  Van Der Kleij, Wilhelmus A.
;    APPLICANT:  Herweijer, Margaretha A.
;    APPLICANT:  Van Beckhoven W.C., Rudolf F.
;    APPLICANT:  Quax, Wilhelmus J.
;    APPLICANT:  Jones, Brian E.
;    TITLE OF INVENTION:  ALKALI-TOLERANT XYLANASES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/501,126
;      FILING DATE:  29-DEC-1995
;      CLASSIFICATION:  425
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  4615-0057.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 396 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-501-126-2

  Query Match             5.8%;  Score 93;  DB 4;  Length 396;
  Best Local Similarity   24.7%;  Pred. No. 0.089;
  Matches   60;  Conservative   36;  Mismatches   83;  Indels   64;  Gaps   15;

Qy        100 VEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPVG--PVTGSLYHLGADKKVKM-HESNI 156
              |||:    :|  :|  | : :   :| :|  ||   |  |     |||| |:   : |:
Db         71 VEPY----QLEGRQAQILKHHYNSLVAENAMKPVSLQPREGEWNWEGADKIVEFARKHNM 126

Qy        157 AIA-NGLAWSNDLKKMYYID-SGKRRVDEYDYDASTLSISNQRPLFTFEKH---EVPGYP 211
              :  : | | : : : ::|| :| | ||| | :      : |  |   | |   | |
Db        127 ELRFHTLVWHSQVPEWFFIDENGNRMVDETDPEKR--KANKQLLLERMENHIKTVVERYK 184

Qy        212 D--------GQTIDEEGNL----WVAVFQGQRIIKI--------------------STQQ 239
              |         : ||: | |    |  :  |   ||:                    :|:
Db        185 DDVTSWDVVNEVIDDGGGLRESEWYQI-TGTDYIKVAFETARKYGGEEAKLYINDYNTEV 243

Qy        240 P---EVLLDTVK------IPDPQV---TSVAFGGPNLDELHV-----TSAGLQLDDSSLD 282
              |   : | : ||       :|   |   : :  | |::::        || ||    : ||
Db        244 PSKRDDLYNLVKDLLEQGVPIDGVGHQSHIQIGWPSIEDTRASFEKFTSLGLDNQVTELD 303

Qy        283 KSL 285
               ||
Db        304 MSL 306


RESULT 4
US-08-282-197C-55
; Sequence 55, Application US/08282197C
; Patent No. 5871730
;  GENERAL INFORMATION:
;    APPLICANT:  Brzezinski, Ryszard
;    APPLICANT:  Dery, Claude V
;    APPLICANT:  Beaulieu, Carole
;    TITLE OF INVENTION:  Thermostable Xylanase DNA, Protein and
;    TITLE OF INVENTION:  Methods of Use
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Ave., NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/282,197C
;      FILING DATE:  29-JUL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cimbala, Michele A
```

```
;      REGISTRATION NUMBER:  33,851
;      REFERENCE/DOCKET NUMBER:  1050.0410000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  397 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
US-08-282-197C-55

  Query Match             5.6%;  Score 91;  DB 2;  Length 397;
  Best Local Similarity   24.7%;  Pred. No. 0.15;
  Matches   60;  Conservative   35;  Mismatches   84;  Indels   64;  Gaps   15;

Qy        100 VEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLP--VGPVTGSLYHLGADKKVKM-HESNI 156
              |||:    :|   :|  | : :   :| :|  | : |  |     ||||  |:   : |:
Db         71 VEPY----QLEGRQAQILKHHYNSLVAENAMKPESLQPREGEWNWEGADKIVEFARKHNM 126

Qy        157 AIA-NGLAWSNDLKKMYYID-SGKRRVDEYDYDASTLSISNQRPLFTFEKH---EVPGYP 211
              :  : | | : :: ::||  | | ||| | |      : |  |   | |     |  |
Db        127 ELRFHTLVWHSQVPEWFFIDEDGNRMVDETDPDKR--EANKQLLLERMENHIKTVVERYK 184

Qy        212 D--------GQTIDEEGNL----WVAVFQGQRIIKI--------------------STQQ 239
              |         : ||: | |    |  :  |   ||:                    :|:
Db        185 DDVTSWDVVNEVIDDGGGLRESEWYQI-TGTDYIKVAFETARKYGGEEAKLYINDYNTEV 243

Qy        240 P---EVLLDTVK------IPDPQV---TSVAFGGPNLDELHV-----TSAGLQLDDSSLD 282
              |   : | : ||      :|   |   : :  | |::::        || ||    : ||
Db        244 PSKRDDLYNLVKDLLEQGVPIDGVGHQSHIQIGWPSIEDTRASFEKFTSLGLDNQVTELD 303

Qy        283 KSL 285
               ||
Db        304 MSL 306


RESULT 5
US-09-397-885-3
; Sequence 3, Application US/09397885
; Patent No. 6323007
; GENERAL INFORMATION:
;  APPLICANT: Moller, Soren
;  APPLICANT: Johansen, Charlotte
;  APPLICANT: Schafer, Thomas
;  APPLICANT: Ostergaard, Peter Rahbek
;  APPLICANT: Hoeck, Lisbeth Hedegaard
;  TITLE OF INVENTION: A 2,6,-b-D-Fructan Hydrolase Enzyme And
;  TITLE OF INVENTION: Process For Using The Enzyme
;  FILE REFERENCE: 5540.200-US
;  CURRENT APPLICATION NUMBER: US/09/397,885
;  CURRENT FILING DATE: 1999-09-17
;  EARLIER APPLICATION NUMBER: PA 1998 01173
;  EARLIER FILING DATE: 1998-09-18
;  EARLIER APPLICATION NUMBER: PA 1998 01623
;  EARLIER FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: 60/101,615
;  EARLIER FILING DATE: 1998-09-24
;  EARLIER APPLICATION NUMBER: 60/111,675
;  EARLIER FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1277
;   TYPE: PRT
;   ORGANISM: Paenibacillus pabuli
US-09-397-885-3

  Query Match             5.6%;  Score 90;  DB 4;  Length 1277;
  Best Local Similarity   21.2%;  Pred. No. 1.3;
  Matches   55;  Conservative   44;  Mismatches  102;  Indels   58;  Gaps   12;

Qy         43 VPSQKKYTFCKVDKL-VSFIIPLAGSPGRFVVSLEREIAILTWD-GVSAAPTSIEAI--- 97
                :|:|:|     :       |  | :   |:  ::: ||::: | | ||    | |  :
Db        664 MPNQRKVMLAWMTNWDYPFEFPTSSWKGQ--LTIPREVSLRTTDEGVRLVQTPITELQKL 721

Qy         98 ---------VNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPVGPVTGSLYHLGADKK 148
                      : | |  ||          ||  |  |   |:| : : |   |:   |   :
Db        722 RHNLYSAQQMTVGPKSKN---------PLEGLTAGAYEIEAEVEI-PANSSVTEFGFQLR 771

Qy        149 VKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVP 208
               :  :        :|:  | : |:              } :|    :   ||| : ||
Db        772 QREGQKTT-----VAYRVDTQNMF------------VDRTTSGDVSFSDLFT-KVHEAS 812

Qy        209 GYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDPQVTSVAF---GGPNLD 265
                |: | :       : :|  :  :::     :|:   |  |||   ::||   ||
Db        813 LKPENQKVK------LRIFVDESSVEVFGNDGKVVFSDVIFPDPAGRAMAFYSLGG---- 862

Qy        266 ELHVTSAGLQLDDSSLDKS 284
              |: |:|  :   |:   ||
Db        863 EVKVSSMKVYALDNIWRKS 881


RESULT 6
US-08-112-630-2
; Sequence 2, Application US/08112630
; Patent No. 5411886
;  GENERAL INFORMATION:
;    APPLICANT:  UDAKA, Shigezo
;    APPLICANT:  SAKAGUCHI, Kenji
;    APPLICANT:  YAMAGATA, Hideo
;    APPLICANT:  DEKKER, Koen
;    TITLE OF INVENTION:  XYLOSEISOMERASE GENE OF THERMUS
;    TITLE OF INVENTION:  AQUATICUS, XYLOSEISOMERASE AND PROCESS FOR PREPARATION OF
;    TITLE OF INVENTION:  FRUCTOSE
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker and Mathis
;      STREET:  George Mason Bldg., Washington & Prince Sts.
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/112,630
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/783,150
;      FILING DATE:  28-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rea, Teresa S.
;      REGISTRATION NUMBER:  30,427
;      REFERENCE/DOCKET NUMBER:  024705-007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;      TELEX:  440 580
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  387 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-112-630-2

  Query Match             5.5%;  Score 89.5;  DB 1;  Length 387;
  Best Local Similarity   19.8%;  Pred. No. 0.21;
  Matches   74;  Conservative   57;  Mismatches  146;  Indels   97;  Gaps   15;

Qy          4 VVEKIAELGKYTVG---------EGPHWDHETQTLYFVDTVEKTFHKYV---------PS 45
```

```
           || |:|||| | |            |  : :     |   :::|  |          |:
Db         38 VVHKLAELGAYGVNLHDEDLIPRGTPPQERDQIVRRFKKALDETGLKVPMVTANLFSDPA 97

Qy         46 QKKYTFCKVDKLVS--------------------FIIPLAGSPGRFVVSLE-----REI 79
              |   |   |  |                        :::     |||    :|     |::
Db         98 FKDGAFTSPDPWVRAYALRKSLETMDLGAELGAEIYVV----WPGREGAEVEATGKARKV 153

Qy         80 AILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPVGPVTGS 139
               || |  |   : |    : :     |     :| |:::  |         || :
Db        154 ----WDWVREALNFMAAYAEDQGYGYRFALEPKPNEPRGDIYFAT--------VGSMLAF 201

Qy        140 LYHLGADKKVKM-----HE--SNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLS 192
              :: |   ::  :     ||  : :   : :| : |  |:::||   :|:  :| |    |
Db        202 IHTLDRPERFGLNPEFAHETMAGLNFVHAVAQALDAGKLFHIDLNDQRMSRFDQDLRFGS 261

Qy        193 ISNQRPLFTFEKHEVPGYPDG--------QTIDEEGNLWVAVFQGQRIIKISTQQPEVLL 244
              : :   |  :  |  ||           :| |||| :|      |   |  :: |
Db        262 ENLKAAFFLVDLLESSGYQGPRHFDAHALRTEDEEG-VWAFARGCMRTYLILKERAEAFR 320

Qy        245 DTVKIPDPQV-----------TSVAFGGPNLDELHVTSAGLQLDDSSLDKSLVNGH-VY 291
              :    ||:|           ::|  ||   |    :  |:  :  |:     |: :
Db        321 E-----DPEVKELLAAYYQEDPAALALLGPYSRE---KAEALKRAELPLEAKRRRGYALE 372

Qy        292 RVTGLGVKGFAGVK 305
              |:  | |:   ||:
Db        373 RLDQLAVEYLLGVR 386


RESULT 7
US-09-134-001C-3945
; Sequence 3945, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3945
;   LENGTH: 1529
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3945

  Query Match             5.5%;  Score 88.5;  DB 4;  Length 1529;
  Best Local Similarity   22.8%;  Pred. No. 2.6;
  Matches   54;  Conservative   29;  Mismatches   87;  Indels   67;  Gaps    9;

Qy         50 TFCKVDKLVSFIIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRL 109
              :||  ||: :        |||: :: :  |   :  ||   |    |           :|
Db        375 SFCNGDKIGALTDRNGLRPGRYTITKDNFIVFSSEVGVIDVPEENVAF--------KGQL 426

Qy        110 NDGK---ADPLGNLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAW-- 164
              | ||    | | |                          | |: :|    ||| | :
Db        427 NPGKLLLVDFLQN-------------------------KVVENNELKTNIANELPYEQ 459

Qy        165 -------SNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTID 217
                    |||  :|| |  .       |:|  ||    ::  | : | :: |     |:
Db        460 WLKDYKNKNDLDNIYYQSS--------DWDDQTLFRLQKQ--FAYTKEDINKYMTDLVIN 509

Qy        218 EEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDPQVTSVAFGGPNLD---ELHVTS 271
              ::  :    :     | :  :|| | :  |   .|||:     | :|   |  |||
Db        510 KKDPIGAMGYDAP--IAVLNDKPESLFNYFKQLFAQVTN-----PPIDAYREKIVTS 559


RESULT 8
US-09-327-536-2
; Sequence 2, Application US/09327536
; Patent No. 6355477
; GENERAL INFORMATION:
;  APPLICANT:  FISCHETTI, Vincent A.
;  APPLICANT:  ROCHA, Claudia
;  TITLE OF INVENTION:  FIBRONECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A
;  TITLE OF. INVENTION:  STREPTOCOCCI
;  FILE REFERENCE:  022927-008
;  CURRENT APPLICATION NUMBER: US/09/327,536
;  CURRENT FILING DATE:  1999-06-08
;  PRIOR APPLICATION NUMBER:  US 08/714,402
;  PRIOR FILING DATE:  1996-09-16
;  NUMBER OF SEQ ID NOS:  2
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1161
;   TYPE: PRT
;   ORGANISM: SFFBP gene
US-09-327-536-2

  Query Match             5.4%;  Score 88;  DB 4;  Length 1161;
  Best Local Similarity   17.5%;  Pred. No. 1.9;
  Matches   64;  Conservative   55;  Mismatches   122;  Indels   124;  Gaps   16;

Qy         13 KYTVGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFV 72
              |:  | |  |  | : | ·              |||                   |  :
Db        780 KFISGSGNEWSFEFKNL----------------KKY---------------NGTGNDI 806

Qy         73 VSLEREIAI-----LTWDGVSAAPTSIEAIVNVEPHIKNNR---LNDGKADPLGNLWTGT 124
              :   :|: :     :|:       |  | |    | ::      |  | :       ||
Db        807 IYSVKEVTVPTGYDVTYSANDIINTKREVITQQGPKLEIEETLPLESGASG-------GT 859

Qy        125 MAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEY 184
              ::   ||  ::|     |    : : | :                    :|   || :|
Db        860 TTVEDSRPVDTLSGLSSEQGQSGDMTIEEDSAT--------------HIKFSKRDIDGK 904

Qy        185 DYDASTLSI--SNQRPLFTF-------EKHEVPG--------YPDGQ--------TIDEE 219
              :   :|: :  |: : : |:       : : :||          |||         |::|:
Db        905 ELAGATMELRDSSGKTISTWISDGQVKDFYLMPGKYTFVETAAPDGYEIATAITFTVNEQ 964

Qy        220 GNLWV---AVFQGQRIIKISTQQP----------EVLLD-------TVKIPDPQVTSVA 258
              | : |   |     ·  |: :   :|          | | |       | :| | : : :
Db        965 GQVTVNGKATKGDTHIVMVDAYKPTKGSGQVIDIEEKLPDEQGHSGSTTEIEDSKSSDLI 1024

Qy        259 FGGPN--LDELHVTSAGL--------QLDDSSLDKSLVNGHVYRV-----TGLGVKGFAG 303
               ||.   :|    | :|:        :::||     ::  |    |     |  |: | :|
Db       1025 IGGQGEVVDTTEDTQSGMTGHSGSTTEIEDSKSSDVIIGGQGQVVETTEDTQTGMYGDSG 1084

Qy        304 VKVKL 308
              | ::
Db       1085 CKTEV 1089


RESULT 9
US-09-548-372D-51
; Sequence 51, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280I
;  CURRENT APPLICATION NUMBER: US/09/548,372D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
```

```
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Hu-Asp2(b) delta TM
US-09-548-372D-51

  Query Match             5.4%;  Score 86.5;  DB 4;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 0.54;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       | |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
              | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||             :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 10
US-09-548-367D-51
; Sequence 51, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280H
;  CURRENT APPLICATION NUMBER: US/09/548,367D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Hu-Asp2(b) delta TM
US-09-548-367D-51

  Query Match             5.4%;  Score 86.5;  DB 4;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 0.54;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       | |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
              | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||             :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 11
US-09-548-372D-53
; Sequence 53, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280I
;  CURRENT APPLICATION NUMBER: US/09/548,372D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
;   LENGTH: 434
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Hu-Asp2(b) delta TM
US-09-548-372D-53

  Query Match             5.4%;  Score 86.5;  DB 4;  Length 434;
  Best Local Similarity   22.0%;  Pred. No. 0.55;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       | |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
              | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||             :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 12
US-09-548-367D-53
; Sequence 53, Application US/09548367D
; Patent No. 6440698
```

```
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280H
;  CURRENT APPLICATION NUMBER: US/09/548,367D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
;   LENGTH: 434
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Hu-Asp2(b) delta TM
US-09-548-367D-53

  Query Match             5.4%;  Score 86.5;  DB 4;  Length 434;
  Best Local Similarity   22.0%;  Pred. No. 0.55;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |       | |   |||  |       |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       | |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |       :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
                | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||             :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 13
US-09-548-372D-6
; Sequence 6, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL..
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280I
;  CURRENT APPLICATION NUMBER: US/09/548,372D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 476
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-548-372D-6

  Query Match             5.4%;  Score 86.5;  DB 4;  Length 476;
  Best Local Similarity   22.0%;  Pred. No. 0.64;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |       | |   |||  |       |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       | |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |       :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
                | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||             :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 14
US-09-548-372D-73
; Sequence 73, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280I
;  CURRENT APPLICATION NUMBER: US/09/548,372D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
;   LENGTH: 476
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-548-372D-73

  Query Match             5.4%;  Score 86.5;  DB 4;  Length 476;
  Best Local Similarity   22.0%;  Pred. No. 0.64;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |       | |   |||  |       |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       | |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |       :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQTEALASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
```

```
              | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||              :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 15
US-09-548-367D-6
; Sequence 6, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280H
;  CURRENT APPLICATION NUMBER: US/09/548,367D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 476
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-548-367D-6

  Query Match             5.4%;  Score 86.5;  DB 4;  Length 476;
  Best Local Similarity   22.0%;  Pred. No. 0.64;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||   |||    |           |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                 |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD-----------EYDYDAST 190
              | |           ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||              :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


Search completed: August  1, 2003, 14:35:38
Job time : 27 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2003, 14:35:09 ; Search time 47 Seconds
(without alignments)
778.256 Million cell updates/sec

Title: US-10-089-986-2
Perfect score: 1615
Sequence: 1 MGPVVEKIAELGKYTVGEGP.....HVYRVTGLGVKGFAGVKVKL 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep1:*
11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep2:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep3:*
13: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	311.5	19.3	286	15	US-10-156-761-13272	Sequence 13272, A
2	92	5.7	4349	15	US-10-160-758-15	Sequence 15, Appl
3	90	5.6	1277	10	US-09-969-362-3	Sequence 3, Appli
4	89.5	5.5	418	15	US-10-156-761-12646	Sequence 12646, A
5	88.5	5.5	610	11	US-09-738-626-4495	Sequence 4495, Ap
6	87	5.4	450	15	US-10-156-761-14443	Sequence 14443, A
7	87	5.4	467	11	US-09-738-626-6671	Sequence 6671, Ap
8	86.5	5.4	428	10	US-09-794-927-51	Sequence 51, Appl
9	86.5	5.4	428	10	US-09-795-847-51	Sequence 51, Appl
10	86.5	5.4	428	10	US-09-794-743-51	Sequence 51, Appl
11	86.5	5.4	428	10	US-09-794-748-51	Sequence 51, Appl
12	86.5	5.4	428	10	US-09-794-925-51	Sequence 51, Appl
13	86.5	5.4	428	10	US-09-681-442-51	Sequence 51, Appl
14	86.5	5.4	428	12	US-09-869-414-51	Sequence 51, Appl
15	86.5	5.4	428	12	US-09-548-366-51	Sequence 51, Appl
16	86.5	5.4	434	10	US-09-794-927-53	Sequence 53, Appl
17	86.5	5.4	434	10	US-09-795-847-53	Sequence 53, Appl
18	86.5	5.4	434	10	US-09-794-743-53	Sequence 53, Appl
19	86.5	5.4	434	10	US-09-794-748-53	Sequence 53, Appl
20	86.5	5.4	434	10	US-09-794-925-53	Sequence 53, Appl
21	86.5	5.4	434	10	US-09-681-442-53	Sequence 53, Appl
22	86.5	5.4	434	12	US-09-869-414-53	Sequence 53, Appl
23	86.5	5.4	434	12	US-09-548-366-53	Sequence 53, Appl
24	86.5	5.4	476	10	US-09-794-927-6	Sequence 6, Appli
25	86.5	5.4	476	10	US-09-795-847-6	Sequence 6, Appli
26	86.5	5.4	476	10	US-09-794-743-6	Sequence 6, Appli
27	86.5	5.4	476	10	US-09-794-748-6	Sequence 6, Appli
28	86.5	5.4	476	10	US-09-794-925-6	Sequence 6, Appli
29	86.5	5.4	476	10	US-09-681-442-6	Sequence 6, Appli
30	86.5	5.4	476	12	US-09-869-414-6	Sequence 6, Appli
31	86.5	5.4	476	12	US-09-548-366-6	Sequence 6, Appli
32	86	5.3	393	14	US-10-080-170-166	Sequence 166, App
33	84.5	5.2	1139	12	US-09-820-843A-15	Sequence 15, Appl
34	83.5	5.2	770	15	US-10-281-478-11	Sequence 11, Appl
35	83	5.1	340	11	US-09-738-626-6960	Sequence 6960, Ap
36	82.5	5.1	462	15	US-10-156-761-11331	Sequence 11331, A
37	82.5	5.1	527	11	US-09-934-868-60	Sequence 60, Appl
38	82.5	5.1	4545	10	US-09-873-403-2	Sequence 2, Appli
39	82	5.1	3472	15	US-10-027-806-4	Sequence 4, Appli
40	82	5.1	3472	15	US-10-034-623-4	Sequence 4, Appli
41	82	5.1	3472	15	US-10-027-801-4	Sequence 4, Appli
42	81.5	5.0	1294	10	US-09-815-242-13724	Sequence 13724, A
43	81.5	5.0	2042	15	US-10-192-584-6	Sequence 6, Appli
44	81	5.0	287	10	US-09-870-962-2	Sequence 2, Appli
45	81	5.0	288	15	US-10-165-696-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-156-761-13272
; Sequence 13272, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13272
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13272

Query Match 19.3%; Score 311.5; DB 15; Length 286;
Best Local Similarity 32.6%; Pred. No. 1.6e-23;
Matches 84; Conservative 45; Mismatches 112; Indels 17; Gaps 8;

```
Qy         16 VGEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSL 75
              :|||| ||   | | :|| :    | | |:  : |    :: |    | || |  ||:|
Db         14 LGEGPTWDTAAQRLIWVDILGSRVHTYDPASGRRTVLATEQHVGAAKPRAG--GGLVVNL 71

Qy         76 EREIAILTWDGVSAAPTSIEAIVNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPVGP 135
                 : :   :|          :: || :  | ||      |:|| |||  |     |
Db         72 RDGVGLYGPEGPGGRDF---RWLHREP-VPGRRGNDAAVAADGSLWAGTMRYDE---AP 123
```

```
Qy        136 VTGSLYHLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISN 195
              |:|  |  |   :    :: |:||: || | : ||: |:  ||:| :| |     :
Db        124 GGGTLSRLAPDGTARTVLDDVTISNGVGWSPDGRLMYFNDTPTRRIDVFDVDEDGRPV-G 182

Qy        196 QRPLFTFEKHEVPGYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTV-KIPDPQV 254
              :| | |  |  |:||| |:| :| :|||:: |  : : :   |   || | ::|  :
Db        183 RRTLATVE--EGAGFPDGLTVDADGCVWVALWDGGVVRRYT---PSGALDRVLELPVRRP 237

Qy        255 TSVAFGGPNLDELHVTSA 272
              |: ||||  | :|::|:|
Db        238 TACAFGGAGLTDLYITTA 255


RESULT 2
US-10-160-758-15
; Sequence 15, Application US/10160758
; Publication No. US20030036076A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: CADs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-089C
;  CURRENT APPLICATION NUMBER: US/10/160,758
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US 60/296,076
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,605
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/357,253
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 4349
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-160-758-15

  Query Match             5.7%;  Score 92;  DB 15;  Length 4349;
  Best Local Similarity   21.9%;  Pred. No. 30;
  Matches   80;  Conservative   41;  Mismatches  140;  Indels  104;  Gaps   16;

Qy         17 GEGPHWDHETQTLYFVDTVEKTFHKYVPSQKKYTFCK----------------------V 54
              |  |||:        |   : :  |:  |:  |||                       :
Db       2666 GGPPHWNSLVPVRLQVVPKKVSLPKF--SEPLYTFSAPEDLPEGSEIGIVKAVAAQDPVI 2723

Qy         55 DKLVSFIIPLAGSPGRFVVSLEREIAIL----TWDGVSAAPTSIEAIVN----------V 100
                ||    | :   |  | ||: :  ::       |  |      |: : :          |
Db       2724 YSLVRGTTPESNKDG--VFSLDPDTGVIKVRKPMDHESTKLYQIDVMAHCLQNTDVVSLV 2781

Qy        101 EPHIKNNRLNDGK----ADPLGNLWTGTMAIDAGLPVGPVT---------GSL-YHLGAD 146
              :|:   :|| :    |||   : |  | :  |  |  ||         | : | | ||
Db       2782 SVNIQVGDVNDNRPVFEADPYKAVLTENMPV--GTSVIQVTAIDKDTGRDGQVSYRLSAD 2839

Qy        147 KKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHE 206
              :||   || :   |   |:::                |  |          |:  |
Db       2840 PGSNVHEL-FAIDSESGWITTLQEL-------------DCETCQ--------TYHFH- 2874

Qy        207 VPGYPDGQTI-------------DEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDP- 252
              |  |  ||||             ||  |      :  |   :   :|  |: |:|  |
Db       2875 VVAYDHGQTIQLSSQALVQVSITDENDNAPRFASEEYRGSVVENSEPGELVATLKTLDAD 2934

Qy        253 ------QVTSVAFGGPNLDELHVTSAGLQLDDSS---LDKSLVNGHVYRVTGLGVKGFAG 303
                   |||      |  | :  ::  | :   ||   ||:     :: |||    |  |
Db       2935 ISEQNRQVTCYITEGDPLGQFGISQVGDEWRISSRKTLDREHTAKYLLRVTASDGKFQAS 2994

Qy        304 VKVKL 308
              | |::
Db       2995 VTVEI 2999


RESULT 3
US-09-969-362-3
; Sequence 3, Application US/09969362
; Patent No. US20020076790A1
; GENERAL INFORMATION:
;  APPLICANT: Moller, Soren
;  APPLICANT:  Johansen, Charlotte
;  APPLICANT:  Schafer, Thomas
;  APPLICANT:  Ostergaard, Peter Rahbek
;  APPLICANT:  Hoeck, Lisbeth Hedegaard
;  TITLE OF INVENTION: A 2,6,-b-D-Fructan Hydrolase Enzyme And
;  TITLE OF INVENTION:  Process For Using The Enzyme
;  FILE REFERENCE: 5540.200-US
;  CURRENT APPLICATION NUMBER: US/09/969,362
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: 09/397,885
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: PA 1998 01623
;  PRIOR FILING DATE: 1998-12-08
;  PRIOR APPLICATION NUMBER: 60/101,615
;  PRIOR FILING DATE: 1998-09-24
;  PRIOR APPLICATION NUMBER: 60/111,675
;  PRIOR FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1277
;   TYPE: PRT
;   ORGANISM: Paenibacillus pabuli
US-09-969-362-3

  Query Match             5.6%;  Score 90;  DB 10;  Length 1277;
  Best Local Similarity   21.2%;  Pred. No. 7.4;
  Matches   55;  Conservative   44;  Mismatches  102;  Indels   58;  Gaps   12;

Qy         43 VPSQKKYTFCKVDKL-VSFIIPLAGSPGRFVVSLEREIAILTWD-GVSAAPTSIEAI--- 97
              :|:|:|     :       |  | :   |:  ::: ||::: | | ||    | |  :
Db        664 MPNQRKVMLAWMTNWDYPFEFPTSSWKGQ--LTIPREVSLRTTDEGVRLVQTPITELQKL 721

Qy         98 ---------VNVEPHIKNNRLNDGKADPLGNLWTGTMAIDAGLPVGPVTGSLYHLGADKK 148
                      : | |  ||          ||  |  |   |:| : : |   |:   |   :
Db        722 RHNLYSAQQMTVGPKSKN---------PLEGLTAGAYEIEAEVEI-PANSSVTEFGFQLR 771

Qy        149 VKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVP 208
               :  :        :|:  | : |:               | :|    :   ||| : ||
Db        772 QREGQKTT-----VAYRVDTQNMF------------VDRTTSGDVSFSDLFT-KVHEAS 812

Qy        209 GYPDGQTIDEEGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDPQVTSVAF---GGPNLD 265
              |: | :       : :|  :  :::      :|:   |  |||   ::||   ||
Db        813 LKPENQKVK------LRIFVDESSVEVFGNDGKVVFSDVIFPDDPAGRAMAFYSLGG---- 862

Qy        266 ELHVTSAGLQLDDSSLDKS 284
              |: |:|  :   |:   ||
Db        863 EVKVSSMKVYALDNIWRKS 881


RESULT 4
US-10-156-761-12646
; Sequence 12646, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
```

```
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12646
;   LENGTH: 418
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-12646

  Query Match             5.5%;  Score 89.5;  DB 15;  Length 418;
  Best Local Similarity   26.5%;  Pred. No. 1.5;
  Matches   26;  Conservative   17;  Mismatches   40;  Indels   15;  Gaps    3;

Qy         44 PSQKKYTFCKVDKLVSFIIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEPH 103
              |::::| :|       |:|| :  |  ||      |   |:: ||| : |:  |:
Db        246 PAERRYPWCLTAMAALLILPLLAA--RLTGSLLALAGALLLAGMATAPTMVTAMTLVQQR 303

Qy        104 IKNNRLNDGKADPLGNLWTGTMAID---AGLPVGPVTG 138
                :|||:|         |:|:     |:  |  ||
Db        304 TPESRLNEGM----------TLAVTGLLGGIACGSATG 331


RESULT 5
US-09-738-626-4495
; Sequence 4495, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4495
;   LENGTH: 610
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-4495

  Query Match             5.5%;  Score 88.5;  DB 11;  Length 610;
  Best Local Similarity   21.7%;  Pred. No. 3.4;
  Matches   70;  Conservative   48;  Mismatches   89;  Indels  115;  Gaps   20;

Qy         14 YTVGEGPHW-----DHETQTLYFVDTVEKTFHKYV-PSQKKYTFCKVDKLVSFIIPLAGS 67
              |  |:   |     | ||:: ||:     |||  : |::              ::  ||
Db        272 YRSGDPEAWRTFWNDPETKSYYFMGKDNITFHSQIWPAE-------------LLGYAGK 317

Qy         68 PGR-------FVVSLEREIA---ILTWDGVSAAPTSIEAIVNVEPHIK------------ 105
                |       |::|  |:     ||  | |   :|   :: |:  :|
Db        318 GSRGGEIGDLGVLNLPTEVVSSEFLTMSG-SKFSSSKGVVIYVKDFLKEFGPDALRYFIA 376

Qy        106 -----NN-----------RLNDGKADPLGNLWTGTMAIDAGLPVG--PVTGSLYHLGADK 147
                  ||           |:|:  |:  |||   |::: |    |  || |:|    :||
Db        377 VAGPENNDTDFTWDEFVRRVNNELANGWGNLVNRTVSM-AHKNFGEVPVPGALEE--SDK 433

Qy        148 KVK-------------------------MH---ESNIAIANGLAWSNDLKKMYYIDSGK 178
              |:                         ||   |:|  ||    |     |:   |: :
Db        434 KILDLATAAFESVAANLDQSKFKAGISEIMHVVGEANAYIAEQEPW-----KLAKDDTKR 488

Qy        179 RRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPD-----GQTIDEEGNLWVAVFQGQRII 233
              |:    : |  : :|:   : |      | |       :|:  :| :| | |   |:
Db        489 ERLATVLWTALQV-VSDCNTMLT------PYLPHTAQKVHETLGRDG-IWAATPQ---IV 537

Qy        234 KISTQQPEVLLDTVKIPDPQVT 255
              ::: : |   :  | :|||: |
Db        538 EVTNESPRQPIG-VGLPDPEHT 558


RESULT 6
US-10-156-761-14443
; Sequence 14443, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14443
;   LENGTH: 450
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-14443

  Query Match             5.4%;  Score 87;  DB 15;  Length 450;
  Best Local Similarity   22.1%;  Pred. No. 3;
  Matches   38;  Conservative   23;  Mismatches   51;  Indels   60;  Gaps    7;

Qy        111 DGKADPLGNLWTGTMAID---AGL----------PVGPVTGSLYHLGADKKVKMHE 153
              ||::|   ::|| | |||   :||              | | :| ::: |||
Db        163 DGQSDGKSSIWTDTFAIDDTASGLRLVSYRLRLTLYRTPGGGITPTVWRLGA-------- 214

Qy        154 SNIAIANGLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQR--PLFTFEKH--EVPG 209
                                  |    | :   ||   :: :   | :: | |  : |
Db        215 ----------------------MGSDIPDRFTVPASAPGLAQELRVPRYSQEIHAGQYPE 252

Qy        210 YPDGQTIDEEGNLWVAVFQGQRIIKISTQQP---EVLLDTVKIPDPQVTSVA 258
              | :|      |  | :    | ||:   ::|   ::       ||||   |
Db        253 YDNG------GEAWCSPTSSQMIIEYWGRRPTPGQLAWVDPSYADPQVCHAA 298


RESULT 7
US-09-738-626-6671
; Sequence 6671, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
```

```
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6671
;   LENGTH: 467
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6671

  Query Match             5.4%;  Score 87;  DB 11;  Length 467;
  Best Local Similarity   35.8%;  Pred. No. 3.2;
  Matches   29;  Conservative   10;  Mismatches   26;  Indels   16;  Gaps    5;

Qy         81 ILTWDGVSAAPTSIEAIVNVEPHI--KNNRLNDGKADPLGNLWTGTMAIDAGLPVGPVTG 138
              | :|  :   || || | |: ||    ||   ||| | :|:     | ||      | ||
Db        391 IRSW--LDGGPTDIENITNICPHDHGNNNDQRDGK-DNMGH-----MNID------PTTG 436

Qy        139 SLYHLGADKKVKMHESNIAIA 159
              : :  ||::  |  :| | |
Db        437 RVGYQPADRRKPMRFNNTAAA 457


RESULT 8
US-09-794-927-51
; Sequence 51, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280FG
;  CURRENT APPLICATION NUMBER: US/09/794,927
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-794-927-51

  Query Match             5.4%;  Score 86.5;  DB 10;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |  :||:     :    ::|   | :

Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
              :|                  |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
              | |          ::  |  |   ::: :| :   ||:           ||:|| |
Db        212 IGGIDH--------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||           :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 9
US-09-795-847-51
; Sequence 51, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280DE
;  CURRENT APPLICATION NUMBER: US/09/795,847
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-795-847-51

  Query Match             5.4%;  Score 86.5;  DB 10;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |  :||:     :    ::|   | :

Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
              :|                  |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD----------EYDYDAST 190
              | |          ::  |  |   ::: :| :   ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
```

```
              :    :| | |   ||              :|||  : |:   | | .
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 10
US-09-794-743-51
; Sequence 51, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280BC
;  CURRENT APPLICATION NUMBER: US/09/794,743
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-794-743-51

  Query Match             5.4%;  Score 86.5;  DB 10;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||  |||   |           |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                  |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD-----------EYDYDAST 190
              | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||              :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 11
US-09-794-748-51
; Sequence 51, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280JL
;  CURRENT APPLICATION NUMBER: US/09/794,748
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-794-748-51

  Query Match             5.4%;  Score 86.5;  DB 10;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |      | |   |||  |      |||  |||   |           |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                  |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD-----------EYDYDAST 190
              | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||              :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 12
US-09-794-925-51
; Sequence 51, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND U
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280HI
;  CURRENT APPLICATION NUMBER: US/09/794,925
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
```

```
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-794-925-51

  Query Match             5.4%;  Score 86.5;  DB 10;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |       | |   |||  |       |||   |||    |           |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |.  |   |:  : | |   :|       |  |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD-----------EYDYDAST 190
                | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||             :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 13
US-09-681-442-51
; Sequence 51, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
;  APPLICANT: Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 28341/6280FG
;  CURRENT APPLICATION NUMBER: US/09/681,442
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/416,901
;  PRIOR FILING DATE: 1999-10-13
;  PRIOR APPLICATION NUMBER: 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-681-442-51

  Query Match             5.4%;  Score 86.5;  DB 10;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |       | |   |||  |       |||   |||    |           |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |. :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
               :|                |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD-----------EYDYDAST 190
                | |          ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||             :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 14
US-09-869-414-51
; Sequence 51, Application US/09869414
; Publication No. US20030077226A1
; GENERAL INFORMATION:
;  APPLICANT: Beinkowski et al.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 28341/6280M
;  CURRENT APPLICATION NUMBER: US/09/869,414
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/416,901
;  PRIOR FILING DATE: 1999-10-13
;  PRIOR APPLICATION NUMBER: 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-869-414-51

  Query Match             5.4%;  Score 86.5;  DB 12;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |       | |   |||  |       |||   |||    |           |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |  :||:     :     ::|   | :
```

```
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
              :|                  |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD-----------EYDYDAST 190
              | |           ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||              :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308


RESULT 15
US-09-548-366-51
; Sequence 51, Application US/09548366
; Publication No. US20030104365A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES THEREFOR
;  FILE REFERENCE:  28341/6280A
;  CURRENT APPLICATION NUMBER: US/09/548,366
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  65
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Hu-Asp2(b)
;   OTHER INFORMATION: delta TM
US-09-548-366-51

  Query Match             5.4%;  Score 86.5;  DB 12;  Length 428;
  Best Local Similarity   22.0%;  Pred. No. 3.2;
  Matches   63;  Conservative   31;  Mismatches   95;  Indels   97;  Gaps   15;

Qy          2 GPVVEKIAEL------GKY---TVGEGPHWDHETQTL-YFVDTVEKTF----------HK 41
              |  || :  |       | |   |||  |      |||   |||    |          |:
Db         58 GSFVEMVDNLRGKSGQGYYVEMTVGSPP------QTLNILVDTGSSNFAVGAAPHPFLHR 111

Qy         42 YVPSQKKYTFCKVDKLVSFIIPLA-----GSPGRFVVSLEREIAILTWDGVSAAPTSIEA 96
              |   |   |:  : | |   :|       |  |  :||:     :     ::|   | :
Db        112 YYQRQLSSTYRDLRKGV--YVPYTQGKWEGELGTDLVSIPHGPNVTVRANIAAITESDKF 169

Qy         97 IVNVEPHIKNNRLNDGKADPLGNLWTGTMAI---------DAGLP------VGPVTGSLY 141
              :|                  |: | | : :          || |      :  | ||:
Db        170 FIN-----------------GSNWEGILGLAYAEIARLCGAGFPLNQSEVLASVGGSMI 211

Qy        142 HLGADKKVKMHESNIAIANGLAWSNDLKKMYYIDSGKRRVD-----------EYDYDAST 190
              | |           ::  |  |   ::: :| :    ||:           ||:|| |
Db        212 IGGIDH---------SLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSI 262

Qy        191 L--SISNQR-PLFTFE--------KHEVPGYPDGQTIDEEGNLWVA 225
              :    :| | |   ||              :|||  : |:   | |
Db        263 VDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLVCWQA 308
```

Search completed: August 1, 2003, 14:43:14
Job time : 50 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:11:42 ; Search time 63 Seconds
(without alignments)
4497.926 Million cell updates/sec

Title: US-10-089-986-1
Perfect score: 924
Sequence: 1 atggggccagttgttgaaaa.....cgggagttaaagtgaagcta 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	43.4	4.7	7218	1	US-08-232-463-14	Sequence 14, Appl
	2	42.2	4.6	4739	3	US-08-685-871-1	Sequence 1, Appli
	3	35.8	3.9	1114	4	US-09-378-088A-85	Sequence 85, Appl
	4	35.8	3.9	1152	3	US-08-844-188-42	Sequence 42, Appl
	5	35.8	3.9	1241	4	US-09-378-088A-42	Sequence 42, Appl
	6	35.8	3.9	2132	3	US-08-844-188-39	Sequence 39, Appl
	7	35.8	3.9	2132	4	US-09-378-088A-39	Sequence 39, Appl
	8	34.8	3.8	162450	4	US-09-345-882-1	Sequence 1, Appli
	9	33.6	3.6	2169	4	US-09-434-408-3	Sequence 3, Appli
c	10	33.6	3.6	43360	4	US-09-453-702B-206	Sequence 206, App
c	11	33.6	3.6	45325	4	US-09-453-702B-261	Sequence 261, App
	12	33.2	3.6	1052	4	US-08-961-527-244	Sequence 244, App
	13	32.8	3.5	1152	3	US-08-844-188-37	Sequence 37, Appl
	14	32.8	3.5	1152	4	US-09-378-088A-37	Sequence 37, Appl
	15	32.8	3.5	2230	3	US-08-844-188-34	Sequence 34, Appl
	16	32.8	3.5	2230	4	US-09-378-088A-34	Sequence 34, Appl
	17	32.6	3.5	1103	4	US-09-378-088A-57	Sequence 57, Appl
	18	32.6	3.5	14507	3	US-08-785-150-1	Sequence 1, Appli
	19	32.6	3.5	14507	4	US-09-660-299-1	Sequence 1, Appli
	20	32.6	3.5	14507	4	US-09-435-377-1	Sequence 1, Appli
c	21	32.4	3.5	270	4	US-09-134-001C-1648	Sequence 1648, Ap
	22	32.4	3.5	13425	4	US-08-961-527-151	Sequence 151, App
	23	32.2	3.5	3312	4	US-09-060-410-1	Sequence 1, Appli
	24	32.2	3.5	7760	4	US-08-961-527-63	Sequence 63, Appl
	25	32	3.5	3667	4	US-09-221-017B-727	Sequence 727, App
	26	31.8	3.4	1725	1	US-08-257-073-14	Sequence 14, Appl
	27	31.6	3.4	624	4	US-08-235-836C-35	Sequence 35, Appl
	28	31.6	3.4	1033	4	US-08-936-165A-7	Sequence 7, Appli
c	29	31.6	3.4	8378	5	PCT-US91-09055-1	Sequence 1, Appli
c	30	31.6	3.4	10660	2	US-08-267-803B-8	Sequence 8, Appli
c	31	31.6	3.4	10660	4	US-09-041-886-16	Sequence 16, Appl
	32	31.4	3.4	426	4	US-09-328-111-304	Sequence 304, App
	33	31.4	3.4	834	3	US-08-633-993A-12	Sequence 12, Appl
	34	31.4	3.4	834	3	US-08-844-188-12	Sequence 12, Appl
	35	31.4	3.4	834	4	US-09-378-088A-12	Sequence 12, Appl
c	36	31.4	3.4	2313	1	US-08-232-538-5	Sequence 5, Appli
c	37	31.4	3.4	2313	4	US-09-427-353-1	Sequence 1, Appli
c	38	31.4	3.4	2352	1	US-08-232-538-17	Sequence 17, Appl
c	39	31.4	3.4	2352	2	US-08-786-164-17	Sequence 17, Appl
c	40	31.4	3.4	2523	4	US-09-051-363-1	Sequence 1, Appli
c	41	31.4	3.4	2651	2	US-08-786-164-5	Sequence 5, Appli
	42	31.4	3.4	5319	1	US-08-169-927-1	Sequence 1, Appli
c	43	31.4	3.4	7032	2	US-08-149-097D-24	Sequence 24, Appl
c	44	31.4	3.4	7032	3	US-08-949-386-24	Sequence 24, Appl
c	45	31.4	3.4	7032	3	US-08-450-562-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 4.7%; Score 43.4; DB 1; Length 7218;

```
  Best Local Similarity   4.3%;  Pred. No. 0.0039;
  Matches   17;  Conservative  211;  Mismatches  167;  Indels    0;  Gaps    0;

Qy        268 GCACCTACAAGCATAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTC 327
              | |  || |   |||||||   ||| ||  :  :::  : : :  ::::: :: :::
Db       1463 GTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1404

Qy        328 AATGATGGCAAAGCAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCT 387
              :: :: :: :::: :::       :: ::   : ::: ::: : :: ::  :  :: :
Db       1403 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1344

Qy        388 GGTCTCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTA 447
              ::       : ::::  ::  :  :: ::   : :  :   :::::  :: :::::::: :
Db       1343 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1284

Qy        448 AAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAG 507
              :::: : : ::::: :: : ::  : :: ::: :::   : : :::: :: ::   ::::
Db       1283 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1224

Qy        508 AAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCT 567
              :::: : :  : :  ::   :::::::::::::: ::: ::: : ::  : :: :
Db       1223 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1164

Qy        568 ACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGA 627
              : :  :   :  :: ::  :: ::  :  :   :     :::::: : :::: :   :::
Db       1163 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1104

Qy        628 TATCCAGATGGTCAAACAATTGATGAGGAGGGTAA 662
               :   ::: ::  ::: ::  :: ::::::::: ::
Db       1103 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1069


RESULT 2
US-08-685-871-1
; Sequence 1, Application US/08685871
; Patent No. 6013499
;  GENERAL INFORMATION:
;    APPLICANT:  NARUMIYA, Shuh
;    APPLICANT:  IWAMATSU, Akihiro
;    TITLE OF INVENTION:  RHO TARGET PROTEIN KINASE P160
;    NUMBER OF SEQUENCES:  68
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,871
;      FILING DATE:  24-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-184102
;      FILING DATE:  25-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-262553
;      FILING DATE:  14-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16887/845
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4739 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  448..4509
US-08-685-871-1

  Query Match             4.6%;  Score 42.2;  DB 3;  Length 4739;
  Best Local Similarity   47.0%;  Pred. No. 0.0076;
  Matches  131;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy        464 ACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTG 523
              |||  |    | ||  ||  | ||  |  |||| |||  |  | |  |   ||| |  ||
Db       1496 ACACTGTAGCACCAGTTGTACCCGATTTAAGTAGTGACATTGATACTAGTAATTTTGATG 1555

Qy        524 ATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCA 583
              | | ||   ||   | || ||| ||  |    |    | | ||| |   | | |  |||
Db       1556 ACTTGGAAGAAGATAAAGGAGAGGAAGAAACATTCCCTATTCCTAAAGCTTTCGTTGGCA 1615

Qy        584 ATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAA 643
              ||||||  || ||  |    |||  | |  ||      | | ||||   |  |     ||
Db       1616 ATCAACTACCTTTTGTAGGATTTACATATTATAGCAATCGTAGATACTTATCTTCAGCAA 1675

Qy        644 CAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAA 703
                | ||||  |  | | |   |    ||| | |          |||  || | |  |||
Db       1676 ATCCTAATGATAACAGAACTAGCTCCAATGCAGATAAAAGCTTGCAGGAAAGTTTGCAAA 1735

Qy        704 TCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAA 742
                |  | | |  | | |||||     ||| |||  | ||
Db       1736 AAACAATCTATAAGCTGGAAGAACAGCTGCATAATGAAA 1774


RESULT 3
US-09-378-088A-85
; Sequence 85, Application US/09378088A
; Patent No. 6372480
; GENERAL INFORMATION:
;  APPLICANT:  Narva, Kenneth E.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Knuth, Mark
;  APPLICANT:  Pollard, Michael R.
;  APPLICANT:  Cardineau, Guy A.
;  APPLICANT:  Schwab, George E.
;  APPLICANT:  Michaels, Tracy E.
;  APPLICANT:  Finstad Lee, Stacy
;  APPLICANT:  Burmeister, Paula
;  APPLICANT:  Dojillo, Joanna
;  TITLE OF INVENTION:  Pesticidal Proteins
;  FILE REFERENCE:  MA703C2
;  CURRENT APPLICATION NUMBER: US/09/378,088A
;  CURRENT FILING DATE:  1999-08-20
;  PRIOR APPLICATION NUMBER:  US 08/633,993
;  PRIOR FILING DATE:  1996-04-19
;  PRIOR APPLICATION NUMBER:  US 08/844,188
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  130
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 85
;   LENGTH: 1114
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-09-378-088A-85

  Query Match             3.9%;  Score 35.8;  DB 4;  Length 1114;
  Best Local Similarity   45.5%;  Pred. No. 0.34;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||   ||  |  |  |   ||| ||| || |    | || |   ||    ||   ||
Db        618 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 677
```

```
Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
              |||   |||  |  ||| | ||| | |  || ||   |  |         |     | |
Db        678 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 737

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
                |  ||    | ||||  |    ||  | | |  ||||| |  |  |||||   |||
Db        738 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 797

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |     | |   | |||   | |    |   || |
Db        798 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 857

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725
              |     ||| ||  |||    |    ||||| | ||| |
Db        858 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 896


RESULT 4
US-08-844-188-42
; Sequence 42, Application US/08844188
; Patent No. 6127180
;  GENERAL INFORMATION:
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Schnepf, H. Ernest
;    APPLICANT:  Knuth, Mark
;    APPLICANT:  Pollard, Michael R.
;    APPLICANT:  Cardineau, Guy
;    APPLICANT:  Schwab, George E.
;    TITLE OF INVENTION:  Pesticidal Toxins
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606-6669
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/844,188
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/633,993
;      FILING DATE:  19-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sanders, Jay M.
;      REGISTRATION NUMBER:  39,355
;      REFERENCE/DOCKET NUMBER:  MA-703C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-844-188-42

  Query Match             3.9%;  Score 35.8;  DB 3;  Length 1152;
  Best Local Similarity   45.5%;  Pred. No. 0.35;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||   ||  |  |  |   |||  ||| || |    | || |   ||    ||   ||
Db        618 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 677
```

```
Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
              |||   |||  |  ||| | ||| | |  || ||   |  |         |     | |
Db        678 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 737

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
                |  ||    | ||||  |    ||  | | |  ||||| |  |  |||||   |||
Db        738 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 797

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |     | |   | |||   | |    |   || |
Db        798 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 857

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725
              |     ||| ||  |||    |    ||||| | ||| |
Db        858 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 896


RESULT 5
US-09-378-088A-42
; Sequence 42, Application US/09378088A
; Patent No. 6372480
; GENERAL INFORMATION:
;  APPLICANT:  Narva, Kenneth E.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Knuth, Mark
;  APPLICANT:  Pollard, Michael R.
;  APPLICANT:  Cardineau, Guy A.
;  APPLICANT:  Schwab, George E.
;  APPLICANT:  Michaels, Tracy E.
;  APPLICANT:  Finstad Lee, Stacy
;  APPLICANT:  Burmeister, Paula
;  APPLICANT:  Dojillo, Joanna
;  TITLE OF INVENTION:  Pesticidal Proteins
;  FILE REFERENCE:  MA703C2
;  CURRENT APPLICATION NUMBER: US/09/378,088A
;  CURRENT FILING DATE:  1999-08-20
;  PRIOR APPLICATION NUMBER:  US 08/633,993
;  PRIOR FILING DATE:  1996-04-19
;  PRIOR APPLICATION NUMBER:  US 08/844,188
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  130
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 1241
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (18)
;   OTHER INFORMATION: Undetermined
;   NAME/KEY: misc_feature
;   LOCATION: (53)
;   OTHER INFORMATION: Undetermined
;   NAME/KEY: misc_feature
;   LOCATION: (61)
;   OTHER INFORMATION: Undetermined
;   NAME/KEY: misc_feature
;   LOCATION: (68)
;   OTHER INFORMATION: Undetermined
;   NAME/KEY: misc_feature
;   LOCATION: (73)
;   OTHER INFORMATION: Undetermined
;   NAME/KEY: misc_feature
;   LOCATION: (81)
;   OTHER INFORMATION: Undetermined
US-09-378-088A-42

  Query Match             3.9%;  Score 35.8;  DB 4;  Length 1241;
  Best Local Similarity   45.5%;  Pred. No. 0.36;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
```

```
Db        705 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 764

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566

Db        765 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 824

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626

Db        825 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 884

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686

Db        885 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 944

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725

Db        945 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 983


RESULT 6
US-08-844-188-39
; Sequence 39, Application US/08844188
; Patent No. 6127180
;  GENERAL INFORMATION:
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Schnepf, H. Ernest
;    APPLICANT:  Knuth, Mark
;    APPLICANT:  Pollard, Michael R.
;    APPLICANT:  Cardineau, Guy
;    APPLICANT:  Schwab, George E.
;    TITLE OF INVENTION:  Pesticidal Toxins
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606-6669
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/844,188
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/633,993
;      FILING DATE:  19-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sanders, Jay M.
;      REGISTRATION NUMBER:  39,355
;      REFERENCE/DOCKET NUMBER:  MA-703C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2132 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-844-188-39

  Query Match             3.9%;  Score 35.8;  DB 3;  Length 2132;
  Best Local Similarity   45.5%;  Pred. No. 0.45;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506

Db       1401 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 1460

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566

Db       1461 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 1520

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626

Db       1521 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 1580

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686

Db       1581 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 1640

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725

Db       1641 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 1679


RESULT 7
US-09-378-088A-39
; Sequence 39, Application US/09378088A
; Patent No. 6372480
; GENERAL INFORMATION:
;  APPLICANT:  Narva, Kenneth E.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Knuth, Mark
;  APPLICANT:  Pollard, Michael R.
;  APPLICANT:  Cardineau, Guy A.
;  APPLICANT:  Schwab, George E.
;  APPLICANT:  Michaels, Tracy E.
;  APPLICANT:  Finstad Lee, Stacy
;  APPLICANT:  Burmeister, Paula
;  APPLICANT:  Dojillo, Joanna
;  TITLE OF INVENTION:  Pesticidal Proteins
;  FILE REFERENCE:  MA703C2
;  CURRENT APPLICATION NUMBER: US/09/378,088A
;  CURRENT FILING DATE:  1999-08-20
;  PRIOR APPLICATION NUMBER:  US 08/633,993
;  PRIOR FILING DATE:  1996-04-19
;  PRIOR APPLICATION NUMBER:  US 08/844,188
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  130
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 2132
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-09-378-088A-39

  Query Match             3.9%;  Score 35.8;  DB 4;  Length 2132;
  Best Local Similarity   45.5%;  Pred. No. 0.45;
  Matches  127;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506

Db       1401 AAAATATCAATATTGGCAACGAGCAGTAGGAAGTAATGTAGCTTTACGTCCACATGAAAA 1460

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566

Db       1461 AAAATCATATACTTATGAATGGGGCACAGAAATAGATCAAAAAACAACAATTATAAATAC 1520

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626

Db       1521 ATTAGGATTTCAAATCAATATAGATTCAGGAATGAAATTTGATATACCAGAAGTAGGTGG 1580

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686

Db       1581 AGGTACAGATGAAATAAAAACACAACTAAATGAAGAATTAAAAATAGAATATAGTCATGA 1640

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGT 725
```

```
                |     ||| ||  |||    |    ||||| | ||| |
Db       1641 AACTAAAATAATGGAAAAATATCAAGAACAATCTGAAAT 1679


RESULT 8
US-09-345-882-1
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
;  APPLICANT: Bougueleret, Lydie
;  TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
;  TITLE OF INVENTION:  AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
;  FILE REFERENCE: GENSET.031A
;  CURRENT APPLICATION NUMBER: US/09/345,882
;  CURRENT FILING DATE:  1999-06-30
;  PRIOR APPLICATION NUMBER: US 60/091,315
;  PRIOR FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: US 60/111,909
;  PRIOR FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 140
;  SOFTWARE: Patent.pm
; SEQ ID NO 1
;   LENGTH: 162450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 72794
;   OTHER INFORMATION: 5-124-273  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 88073
;   OTHER INFORMATION: 5-127-261  :  polymorphic  base  A  or  C
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 90842
;   OTHER INFORMATION: 99-1437-325  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 93714
;   OTHER INFORMATION: 5-128-60  :  polymorphic  base  deletion  of  GT
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97122
;   OTHER INFORMATION: 99-1442-224  :  polymorphic  base  G  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97152
;   OTHER INFORMATION: 5-129-144  :  polymorphic  base  deletion  of  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99098
;   OTHER INFORMATION: 5-130-257  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99117
;   OTHER INFORMATION: 5-130-276  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 103806
;   OTHER INFORMATION: 5-131-395  :  polymorphic  base  A  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 106940
;   OTHER INFORMATION: 5-133-375  :  polymorphic  base  insertion  of  A
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108106
;   OTHER INFORMATION: 5-135-155  :  polymorphic  base  insertion  of  A
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108149
;   OTHER INFORMATION: 5-135-198  :  polymorphic  base  insertion  of  GTTT
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108308
;   OTHER INFORMATION: 5-135-357  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108471
;   OTHER INFORMATION: 5-136-174  :  polymorphic  base  C  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 134134
;   OTHER INFORMATION: 5-140-120  :  polymorphic  base  C  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 134362
;   OTHER INFORMATION: 5-140-348  :  polymorphic  base  insertion  of  A
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 134374
;   OTHER INFORMATION: 5-140-361  :  polymorphic  base  insertion  of  CA
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 146328
;   OTHER INFORMATION: 5-143-84  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 146345
;   OTHER INFORMATION: 5-143-101  :  polymorphic  base  A  or  C
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 150329
;   OTHER INFORMATION: 5-145-24  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 160031
;   OTHER INFORMATION: 5-148-352  :  polymorphic  base  G  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 72771..72817
;   OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 72771..72817
;   OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 88050..88096
;   OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 88050..88096
;   OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 90819..90865
;   OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 90819..90865
;   OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 93690..93736
;   OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 93690..93736
;   OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97099..97145
;   OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
;   FEATURE:
```

```
;   NAME/KEY: allele
;   LOCATION: 97099..97145
;   OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97130..97177
;   OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97130..97177
;   OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99075..99121
;   OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99075..99121
;   OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99094..99140
;   OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99094..99140
;   OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 103783..103828
;   OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 103783..103828
;   OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 106918..106966
;   OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 106918..106966
;   OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108084..108130
;   OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108084..108130
;   OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108127..108177
;   OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108127..108177
;   OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
;   FEATURE:

  Query Match             3.8%;  Score 34.8;  DB 4;  Length 162450;
  Best Local Similarity   51.3%;  Pred. No. 5.3;
  Matches   81;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy         38 AGTATACGGTTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCG 97
              ||  ||   ||  ||||| | | || |||| |||| ||   |      |||     |
Db      57901 AGCTTAAATTTTAAGAAGATACACAGTGGGGCCATGTAAAAAACCAAAAATAAACTTTAA 57960

Qy         98 ACACCGTAGAGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTA 157
              | |      |||| | | || | ||||||  |     ||  || | | |  ||||||
Db      57961 AAAATTGTAAGAAGAATATTAAGAAATATAGATAACATCCAAAGATTTTCGTGTTTTGGG 58020

Qy        158 AAGTAGATAAACTGGTTTCTTTCATTATTCCCCTTGCT 195
              |||  | | |  |  ||| |||  || |     ||| |
Db      58021 AAGGGGTTGAGTTTTTTTGTTTTGTTTTGTTTTTTGTT 58058


RESULT 9
US-09-434-408-3
; Sequence 3, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
;  APPLICANT: Venezia, Domenick
;  APPLICANT: Grossmann, Angelika
;  TITLE OF INVENTION: RING FINGER PROTEIN ZAPOP3
;  FILE REFERENCE: 98-41
;  CURRENT APPLICATION NUMBER: US/09/434,408
;  CURRENT FILING DATE: 1999-11-04
;  EARLIER APPLICATION NUMBER: US 60/108,258
;  EARLIER FILING DATE: 1998-11-12
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 2169
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Degenerate polynucleotide sequence of zapop3
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2169)
;   OTHER INFORMATION: n = A,T,C or G
US-09-434-408-3

  Query Match             3.6%;  Score 33.6;  DB 4;  Length 2169;
  Best Local Similarity   21.5%;  Pred. No. 2;
  Matches  113;  Conservative  115;  Mismatches  290;  Indels    7;  Gaps    1;

Qy        136 CAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTCATTATTCCCCTTGCT 195
              | |      ||::     || :   |:| : |::: || :|     |:||: | || ::
Db        590 CNGCNGCNATHYTNCARTTYYTNTGYAARGARWSNGGNYTNGARTAYTAYCCNCCNWSNC 649

Qy        196 GGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCCATTCTTACATGGGAT 255
              :  ::  :  ||  |::     ||:  :|| |    :||::: :    ::       |
Db        650 ARTAYYTNYTNCCNATHYTNGARCARGAYGGNATHGARAAYWSNMGNGAYWSNCCNGAYG 709

Qy        256 GGCGTTAGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTCGAACCACACATTAAA 315
              | |       : |   ::  :   |   :    | ||  |:  ::|   :::  |:|| |
Db        710 GNCCNACNGAYMGNTTYWSNMGNGARGARYTNGARTGGCARAAYMGNTTYWSNGAYTAYG 769

Qy        316 AATAAC-------AGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTATGGACAGG 368
              |: |::       : |   :||| | | :||::       :  :   :  :  | ::  |
Db        770 ARAARMGNAARGARCARAARATGYTNGARAARYTNGARTTYGARMGNMGNYTNGARYTNG 829

Qy        369 TACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTT 428
                ||: |   :   || | : | :   | : |:     : :: | :  :| : | ||: |
Db        830 GNCARMGNGARCAYACNCARYTNYTNCARCARWSNWSNWSNCARAARGAYGARATHYTNC 889

Qy        429 AGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGC 488
              |:    | | : |:  : |::   |       :  :   :       |:||   :  |  :
Db        890 ARACNGTNAARGARGARCARWSNMGNYTNGARCARGGNYTNWSNGARCAYCARMGNCAYY 949

Qy        489 GTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGA 548
              | |    ::    ::|            ::| |  | :    |: |: |    :: :
Db        950 TNGAYGCNGARMGNCARMGNYTNCARGARCARYTNAARCARACNGARCARAAYATHWSNW 1009

Qy        549 GTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGA 608
              : :  ||:  :| ::| :| ||  |:     :      ||       :    : |::| |
Db       1010 SNMGNATHCARAARYTNYTNCARGAYAAYCARMGNCARAARAARWSNWSNGARATHYTNA 1069

Qy        609 AAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGA 653
              |::  |  |:  :  :   ||::  ||||   |::    ||: |
Db       1070 ARWSNYTNGARAAYGARMGNATHMGNATGGARCARYTNATGWSNA 1114
```

```
RESULT 10
US-09-453-702B-206/c
; Sequence 206, Application US/09453702B
; Patent No. 6365723
;  GENERAL INFORMATION:
;       APPLICANT: Blattner, Frederick R.
;                  Burland,     Valerie
;                  Perna,  Nicole T.
;                  Plunkett, Guy
;                  Welch,  Rod
;       TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Quarles & Brady
;            STREET: 1 South Pinckney Street
;            CITY: Madison
;            STATE: WI
;            COUNTRY: US
;            ZIP: 53701-2113
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Word Perfect 8.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/453,702B
;            FILING DATE: 03-Dec-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/110,955
;            FILING DATE: 04-DEC-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Seay, Nicholas J.
;            REGISTRATION NUMBER: 27386
;            REFERENCE/DOCKET NUMBER: 960296.95017
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (608) 251-5000
;            TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 206:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 43360
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA  (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-09-453-702B-206

  Query Match             3.6%;  Score 33.6;  DB 4;  Length 43360;
  Best Local Similarity   45.7%;  Pred. No. 7.1;
  Matches  117;  Conservative    0;  Mismatches  139;  Indels    0;  Gaps    0;

Qy        455 ACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGT 514
              | |||   |||||   || |    ||     |  ||||   ||  ||  ||| |||
Db      14108 ATGAGGATAACATTTTTAAATTTTATTCATCCTGGTGGTTGAACAATGAGAAAAAATATG 14049

Qy        515 ATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTAT 574
              | |  ||   |||   || |  ||||||||   | |    |   |   | |||  | ||
Db      14048 GTGAACTTACATCGTCTAAGAATAGAGTAGAAAGGAAGCTGTTGTACAGTACTATTTAAT 13989

Qy        575 CCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAG 634
              |  |   || |      ||||  |   |||  || | | | |    || || |
Db      13988 GGAATATTCATAATTAATGATTACGTGAGTTTCCAATGGCTAATGGTATTGAATTTAATC 13929

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
              |    |  ||  || |    |  |||||  ||  | |   |  ||  ||   |||| |
Db      13928 AAAACCCCGCATCTGTTTTTAATTCTAATTCATTAGATTTTGAATTAGAATCTCAGCAAT 13869

Qy        695 TTATTAAAATCAGTAC 710
              | |   |||  | | |
Db      13868 TAACGCAAAAAAATTC 13853
```

```
RESULT 11
US-09-453-702B-261/c
; Sequence 261, Application US/09453702B
; Patent No. 6365723
;  GENERAL INFORMATION:
;       APPLICANT: Blattner, Frederick R.
;                  Burland,     Valerie
;                  Perna,  Nicole T.
;                  Plunkett, Guy
;                  Welch,  Rod
;       TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Quarles & Brady
;            STREET: 1 South Pinckney Street
;            CITY: Madison
;            STATE: WI
;            COUNTRY: US
;            ZIP: 53701-2113
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Word Perfect 8.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/453,702B
;            FILING DATE: 03-Dec-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/110,955
;            FILING DATE: 04-DEC-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Seay, Nicholas J.
;            REGISTRATION NUMBER: 27386
;            REFERENCE/DOCKET NUMBER: 960296.95017
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (608) 251-5000
;            TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 261:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 45325 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 261:
US-09-453-702B-261

  Query Match             3.6%;  Score 33.6;  DB 4;  Length 45325;
  Best Local Similarity   45.7%;  Pred. No. 7.2;
  Matches  117;  Conservative    0;  Mismatches  139;  Indels    0;  Gaps    0;

Qy        455 ACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGT 514
              | |||   |||||   || |    ||     |  ||||   ||  ||  ||| |||
Db      15010 ATGAGGATAACATTTTTAAATTTTATTCATCCTGGTGGTTGAACAATGAGAAAAAATATG 14951

Qy        515 ATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTAT 574
              | |  ||   |||   || |  ||||||||   | |    |   |   | |||  | ||
Db      14950 GTGAACTTACATCGTCTAAGAATAGAGTAGAAAGGAAGCTGTTGTACAGTACTATTTAAT 14891

Qy        575 CCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAG 634
              |  |   || |      ||||  |   |||  || | | | |    || || |
Db      14890 GGAATATTCATAATTAATGATTACGTGAGTTTCCAATGGCTAATGGTATTGAATTTAATC 14831

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
              |    |  ||  || |    |  |||||  ||  | |   |  ||  ||   |||| |
Db      14830 AAAACCCCGCATCTGTTTTTAATTCTAATTCATTAGATTTTGAATTAGAATCTCAGCAAT 14771

Qy        695 TTATTAAAATCAGTAC 710
              | |   |||  | | |
Db      14770 TAACGCAAAAAAATTC 14755
```

```
RESULT 12
US-08-961-527-244
; Sequence 244, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  244:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1052 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-244

  Query Match             3.6%;  Score 33.2;  DB 4;  Length 1052;
  Best Local Similarity   50.6%;  Pred. No. 2;
  Matches   80;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        202 CCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTT 261
              |||||    ||||  |||| |  ||||  | ||||    | | |  || | | .||   |
Db        583 CCTGGTTTCTTTGGTGTCACTAAGGAAAATCAAATCTGTACTTTCTCACGTGGAGGTTCT 642

Qy        262 AGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAAC 321
                | | ||  | | | ||| |  |||  ||| || |  || |   |    ||   ||
Db        643 GATATTACAGGTTCTATCATTGCTGCTGGTGTCAAAGCTGACCTCTATGAAAACTTTACG 702

Qy        322 AGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCT 359
                |  ||| |  |   ||||  | || ||| | | |
Db        703 GACGTTGATGGTATCTTTGCAGCCCACCCTGGTATTAT 740


RESULT 13
US-08-844-188-37
; Sequence 37, Application US/08844188
; Patent No. 6127180
;  GENERAL INFORMATION:
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Schnepf, H. Ernest
;    APPLICANT:  Knuth, Mark
;    APPLICANT:  Pollard, Michael R.
;    APPLICANT:  Cardineau, Guy
;    APPLICANT:  Schwab, George E.
;    TITLE OF INVENTION:  Pesticidal Toxins
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606-6669.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/844,188
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/633,993
;      FILING DATE:  19-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sanders, Jay M.
;      REGISTRATION NUMBER:  39,355
;      REFERENCE/DOCKET NUMBER:  MA-703C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-844-188-37

  Query Match             3.5%;  Score 32.8;  DB 3;  Length 1152;
  Best Local Similarity   44.9%;  Pred. No. 2.7;
  Matches  124;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||   ||  |  |  |   |||  ||| || |    | || |   ||    ||   ||
Db        618 AAAATATCAATATTGGCAACAAGCAGTAGGAAGTAATGTAGCTTTACGTCCGCATGAAAA 677

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
               |||   |||  |  ||| | ||| | |  || ||   |  |       |   | | |
Db        678 AAAATCATATGCTTATGAGTGGGGTACAGAAATAGATCAAAAAACAACTATCATTAATAC 737

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
                |  ||     |  ||| |     |   | | |  ||||| |  |  |||||   |||
Db        738 ATTAGGATTTCAGATTAATATAGATTCGGGAATGAAATTTGATATACCAGAAGTAGGTGG 797

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |     | |   | |||    | |    | ||  |
Db        798 AGGTACAGATGAAATAAAAACACAATTAAACGAAGAATTAAAAATAGAATATAGCCGTGA 857

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGA 722
              |     ||| ||  |||    |    ||||| | ||
Db        858 AACCAAAATAATGGAAAAATATCAGGAACAATCAGA 893


RESULT 14
US-09-378-088A-37
; Sequence 37, Application US/09378088A
; Patent No. 6372480
; GENERAL INFORMATION:
;  APPLICANT:  Narva, Kenneth E.
;  APPLICANT:  Schnepf, H. Ernest
;  APPLICANT:  Knuth, Mark
;  APPLICANT:  Pollard, Michael R.
;  APPLICANT:  Cardineau, Guy A.
;  APPLICANT:  Schwab, George E.
;  APPLICANT:  Michaels, Tracy E.
;  APPLICANT:  Finstad Lee, Stacy
```

```
;  APPLICANT:  Burmeister, Paula
;  APPLICANT:  Dojillo, Joanna
;  TITLE OF INVENTION:  Pesticidal Proteins
;  FILE REFERENCE:  MA703C2
;  CURRENT APPLICATION NUMBER: US/09/378,088A
;  CURRENT FILING DATE:  1999-08-20
;  PRIOR APPLICATION NUMBER:  US 08/633,993
;  PRIOR FILING DATE:  1996-04-19
;  PRIOR APPLICATION NUMBER:  US 08/844,188
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  130
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 37
;   LENGTH: 1152
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-09-378-088A-37

  Query Match             3.5%;  Score 32.8;  DB 4;  Length 1152;
  Best Local Similarity   44.9%;  Pred. No. 2.7;
  Matches  124;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||   ||  |  |  |   ||| ||| || |    | || |   ||    ||   ||
Db        618 AAAATATCAATATTGGCAACAAGCAGTAGGAAGTAATGTAGCTTTACGTCCGCATGAAAA 677

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
               |||   |||  |  ||| | ||| | |  || ||   |  |       |     | | |
Db        678 AAAATCATATGCTTATGAGTGGGGTACAGAAATAGATCAAAAAACAACTATCATTAATAC 737

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
                |  ||     |  |||  |    |   | | |  ||||| |  |  |||||   |||
Db        738 ATTAGGATTTCAGATTAATATAGATTCGGGAATGAAATTTGATATACCAGAAGTAGGTGG 797

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |    | |   | |||    | |    |  ||  |
Db        798 AGGTACAGATGAAATAAAAACACAATTAAACGAAGAATTAAAAATAGAATATAGCCGTGA 857

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGA 722
              |     ||| ||  |||    |    ||||| | ||
Db        858 AACCAAAATAATGGAAAAATATCAGGAACAATCAGA 893


RESULT 15
US-08-844-188-34
; Sequence 34, Application US/08844188
; Patent No. 6127180
;  GENERAL INFORMATION:
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Schnepf, H. Ernest
;    APPLICANT:  Knuth, Mark
;    APPLICANT:  Pollard, Michael R.
;    APPLICANT:  Cardineau, Guy
;    APPLICANT:  Schwab, George E.
;    TITLE OF INVENTION:  Pesticidal Toxins
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606-6669
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/844,188
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/633,993
;      FILING DATE:  19-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sanders, Jay M.
;      REGISTRATION NUMBER:  39,355
;      REFERENCE/DOCKET NUMBER:  MA-703C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2230 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-844-188-34

  Query Match             3.5%;  Score 32.8;  DB 3;  Length 2230;
  Best Local Similarity   44.9%;  Pred. No. 3.6;
  Matches  124;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506
              ||||   ||  |  |  |   ||| ||| || |    | || |   ||    ||   ||
Db       1575 AAAATATCAATATTGGCAACAAGCAGTAGGAAGTAATGTAGCTTTACGTCCGCATGAAAA 1634

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566
               |||   |||  |  ||| | ||| | |  || ||   |  |       |     | | |
Db       1635 AAAATCATATGCTTATGAGTGGGGTACAGAAATAGATCAAAAAACAACTATCATTAATAC 1694

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGG 626
                |  ||     |  |||  |    |   | | |  ||||| |  |  |||||   |||
Db       1695 ATTAGGATTTCAGATTAATATAGATTCGGGAATGAAATTTGATATACCAGAAGTAGGTGG 1754

Qy        627 ATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |  | ||||||    || ||   |    | |   | |||    | |    |  ||  |
Db       1755 AGGTACAGATGAAATAAAAACACAATTAAACGAAGAATTAAAAATAGAATATAGCCGTGA 1814

Qy        687 ACAGCGAATTATTAAAATCAGTACCCAACAACCGGA 722
              |     ||| ||  |||    |    ||||| | ||
Db       1815 AACCAAAATAATGGAAAAATATCAGGAACAATCAGA 1850


Search completed: August  1, 2003, 13:22:13
Job time : 74 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:07:04 ; Search time 1751 Seconds
(without alignments)
8546.333 Million cell updates/sec

Title: US-10-089-986-1
Perfect score: 924
Sequence: 1 atggggccagttgttgaaaa.....cgggagttaaagtgaagcta 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	106.6	11.5	638	13	BI514926	BI514926 BB160016B
2	98.2	10.6	666	10	AV401529	AV401529 AV401529
3	71.2	7.7	670	13	BI611823	BI611823 RH18245.5
4	66.4	7.2	635	13	BI617530	BI617530 RH47894.5
5	65.4	7.1	719	9	AJ425659	AJ425659 AJ425659
6	65.2	7.1	655	13	BI617627	BI617627 RH48022.5
7	65	7.0	560	14	C92366	C92366 C92366 Dict
8	65	7.0	606	14	C92731	C92731 C92731 Dict
9	65	7.0	639	14	C92540	C92540 C92540 Dict
10	65	7.0	679	14	C92417	C92417 C92417 Dict
11	65	7.0	730	14	C90258	C90258 C90258 Dict
c 12	65	7.0	743	13	BJ435579	BJ435579 BJ435579
13	64.4	7.0	555	9	AI663645	AI663645 ui99h04.y
c 14	63.4	6.9	730	10	AV890554	AV890554 AV890554
c 15	62.4	6.8	513	13	BJ371993	BJ371993 BJ371993
c 16	62.4	6.8	582	13	BJ402499	BJ402499 BJ402499
17	62	6.7	592	13	BI161908	BI161908 RE01371.5
c 18	61.4	6.6	742	13	BJ535587	BJ535587 BJ535587
c 19	61.2	6.6	704	10	AV905180	AV905180 AV905180
20	61	6.6	583	13	BI364230	BI364230 RE49294.5
c 21	60.2	6.5	735	13	BJ022574	BJ022574 BJ022574
c 22	59.6	6.5	524	10	AV839733	AV839733 AV839733
23	59	6.4	457	13	BI592225	BI592225 RH09865.5
24	59	6.4	555	13	BI241345	BI241345 RE38383.5
c 25	58.6	6.3	515	12	BG639989	BG639989 SD09890.3
26	57.4	6.2	455	13	BI357410	BI357410 RE43884.5
27	57.4	6.2	478	13	BI615830	BI615830 RH45670.5
28	57.4	6.2	480	13	BI361794	BI361794 RE46157.5
29	57.4	6.2	529	13	BI584695	BI584695 RH23952.5
30	57.4	6.2	531	13	BI364664	BI364664 RE49831.5
31	57.4	6.2	546	13	BI367889	BI367889 RE53758.5
32	57.4	6.2	550	13	BI567910	BI567910 RH39220.5
33	57.4	6.2	551	13	BI586938	BI586938 RH27167.5
34	56.6	6.1	415	9	AA387937	AA387937 vc87h05.r
35	56.2	6.1	617	13	BJ522043	BJ522043 BJ522043
36	56.2	6.1	744	13	BI328880	BI328880 602984879
c 37	56	6.1	613	13	BJ373089	BJ373089 BJ373089
38	55.4	6.0	572	13	BJ488689	BJ488689 BJ488689
39	54.8	5.9	491	13	BI365812	BI365812 RE51194.5
c 40	54.8	5.9	718	13	BJ509252	BJ509252 BJ509252
c 41	54.6	5.9	590	12	BG638003	BG638003 LD12015.3
42	54.4	5.9	508	9	AU242394	AU242394 AU242394
43	53.4	5.8	593	12	BF607009	BF607009 MY2_00006
44	52.6	5.7	774	9	AI047491	AI047491 ud66d11.y
45	52	5.6	440	9	AI785853	AI785853 uj56a07.y

ALIGNMENTS

RESULT 1
BI514926
LOCUS BI514926 638 bp mRNA . linear EST 08-APR-2002
DEFINITION BB160016B20D02.5 Bee Brain Normalized Library, BB16 Apis mellifera cDNA clone BB160016B20D02 5', mRNA sequence.
ACCESSION BI514926
VERSION BI514926.1 GI:15365300
KEYWORDS EST.
SOURCE honeybee.
ORGANISM Apis mellifera
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.
REFERENCE 1 (bases 1 to 638)
AUTHORS Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L., Pardinas,J., Robertson,H.M., Soares,B. and Robinson,G.E.
TITLE Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee
JOURNAL Genome Res. 12 (4), 555-566 (2002)
MEDLINE 21929762
COMMENT Contact: Gene E. Robinson
Department of Entomology
University of Illinois
505 S. Goodwin Ave., Urbana, IL 61801, USA
Tel: 217 265 0309
Fax: 217 244 3499
Email: generobi@life.uiuc.edu
This research was funded by the University of Illinois Critical Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation

Award in Functional Genomics to G.E. Robinson and an NSF
Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.
PCR PRimers
FORWARD: TAATACGACTCACTATAGGG
BACKWARD: ATTAACCCTCACTAAAG
Plate: BB160016B20 row: D column: 02
Seq primer: AGCGGATAACAATTTCACACAGGA
High quality sequence stop: 638.

```
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Apis mellifera"
                     /strain="mixed strains of European bees, predominantly
                     A.m. ligustica"
                     /db_xref="taxon:7460"
                     /clone="BB160016B20D02"
                     /clone_lib="Bee Brain Normalized Library, BB16"
                     /sex="female"
                     /tissue_type="brain"
                     /dev_stage="adult worker honey bee"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3-Pac; Site_1: EcoR1;
                     Site_2: Not1; The BB16 library was contributed by the
                     Soares laboratory and it was constructed and normalized
                     as described by Bonaldo, M.F., Lennon, G. and Soares,
                     M.B. (1996), Genome Research 6(9): 791-806. RNA was
                     prepared from dissected brains of adult worker bees of
                     various ages and various behavioral groups. "
BASE COUNT      223 a     84 c    110 g    221 t
ORIGIN
```

```
  Query Match            11.5%;  Score 106.6;  DB 13;  Length 638;
  Best Local Similarity  50.6%;  Pred. No. 4.2e-19;
  Matches  312;  Conservative    0;  Mismatches  299;  Indels    6;  Gaps    2;

Qy         26 CAGAACTTGGCAAGTATACGGTTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCT 85
              ||  | | ||||  | |      | | ||||||||||| ||||||    | | ||| |
Db         22 CAATAATAGGCAGATTTGATCATAGTGAAGGTCCTCATTGGGATTCACATATTCAAAAAC 81

Qy         86 TATATTTCGTCGACACCGTAGAGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAAT 145
              | ||||| || || |  | ||   |||   |  ||| || ||  |||       ||| |
Db         82 TTTATTTTGTTGATATAGAAGCACAAAAAATATATAGATTTGATCCTGTAACAAAAGATC 141

Qy        146 ACACGTTT---TGTAAAGTAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCC 202
                | | |    | || ||   ||  ||   ||  || |||||||  |||  |||    |
Db        142 TTTCCTGTATATTTATAGAGAATGGACCTATTGGATTTGTTATTCCTGTTGAAGGAGAAC 201

Qy        203 CTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTA 262
              ||      ||||| | | |||  |     ||  |     |  |||| |||||||    ||
Db        202 CTCATAAGTTTGTTGCCGGTTGTGGTACAGATTTCATTTTAGTTACGTGGGATGAACATA 261

Qy        263 GTGCTGCACCTACAAGCAT---AGAAGCTATTGTTAATGTCGAACCACACATTAAAAATA 319
              |  |||  | | |   ||     |||   ||  || ||| ||      |    ||    |
Db        262 GAAATGCTACCAAATCTATTCCTCAAGTACTTAGTATTGTTGATAATGATCGAAATGGAA 321

Qy        320 ACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTA 379
                |||   |||||||| |||||||||  |  | || |   | |||   || |||||  |
Db        322 CAAGATGGAATGATGGAAAAGCAGATTCTTTAGGAAGATTTTGGGGTGGAACAATAGGAC 381

Qy        380 TTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATA 439
              ||      |      ||    ||       |    |||||||| | | |   | |||
Db        382 CAGAAATAAATGATGTTGTTATTCCTAATCAAGCATCTTTATATCGTATTGATTCAGATC 441

Qy        440 AAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATG 499
              ||  |||||   |           |  |||     |||||  | || |||| |  |
Db        442 TTAAACCAAAAAAAGAATTATCTCCTGTTACTAATAGTAATGGATTAGCTTGGAATCTTC 501

Qy        500 ATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATG 559
              |   || | | |||||||||||| |   | |     |||||  |  |  |||| ||
Db        502 AAGATAATACATTTTATTATATTGATACTCCTACACTTCAAGTAGCAGCATTTGATTTTG 561

Qy        560 ATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCATGAAG 619
              |  ||  ||    | | ||    ||| || |      ||||  | || ||||  || |
Db        562 AACCTATTAATGGAACTATATCTAATAAAAGAATTGCATTTGATCTTCAAAAAAATAATA 621
```

```
Qy        620 TGCCTGGATATCCAGAT 636
              |  |||||   || |||
Db        622 TTTCTGGAATACCTGAT 638
```

```
RESULT 2
AV401529
LOCUS       AV401529                 666 bp    mRNA    linear   EST 05-FEB-2000
DEFINITION  AV401529 Bombyx mori diapause-cancelled embryo p50 76 hr after HCl
            treatment (HCl at 20 hr after oviposition) Bombyx mori cDNA clone
            e96h0839 T3, mRNA sequence.
ACCESSION   AV401529
VERSION     AV401529.1  GI:6905181
KEYWORDS    EST.
SOURCE      domestic silkworm.
  ORGANISM  Bombyx mori
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
            Bombycoidea; Bombycidae; Bombyx.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
  TITLE     Bombyx mori cDNA
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Mita K
            Genome Research Group
            National Institute of Radiological Sciences
            Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
            Email: kmita@nirs.go.jp
            method:uni-directional, sequence direction:sequenced from T3 primer
            (5' -> 3')
             Project='Silkworm Genome Program in MAFF, and Research for the
            Future Program in JSPS'. see 'SilkBase',
            <http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.
FEATURES             Location/Qualifiers
     source          1. .666
                     /organism="Bombyx mori"
                     /strain="p50"
                     /db_xref="taxon:7091"
                     /clone="e96h0839"
                     /clone_lib="Bombyx mori diapause-cancelled embryo p50 76
                     hr after HCl treatment (HCl at 20 hr after oviposition)"
                     /sex="female/male mixed"
                     /tissue_type="diapause-cancelled embryo"
                     /dev_stage="76 hr after HCl treatment (HCl at 20 hr after
                     oviposition)"
BASE COUNT      229 a    114 c    154 g    167 t      2 others
ORIGIN
```

```
  Query Match            10.6%;  Score 98.2;  DB 10;  Length 666;
  Best Local Similarity  56.8%;  Pred. No. 1e-16;
  Matches  200;  Conservative    0;  Mismatches  149;  Indels    3;  Gaps    1;

Qy         43 ACGGTTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACC 102
              ||  | |.| ||||| || |||||||||| |    |  ||  |  |||| || ||   || |
Db        210 ACACTGGGGGAAGGGCCCCACTGGGATGAACGNAAGCAAGCGCTATACTTTGTGAGCATC 269

Qy        103 GTAGAGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTA 162
               ||| ||||||  | || ||||||||||||  ||    ||||| | ||       |  ||
Db        270 CAAGATAAAACTATACACAAATATGTACCAACTACTGAAAAACATACAAAAACCAGTTTA 329

Qy        163 GATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGT 222
              |||      ||||  ||||||  | ||  | |  ||  |  | | ||  ||||||||  |
Db        330 GATGGGAGGGTTGGTTTCATATTACCAGTGGAAGGTACAACAGACCAATTTGTAGTGGGA 389

Qy        223 TTGGAACGTGAAATAGCCATTCTTACATGGGATGG---CGTTAGTGCTGCACCTACAAGC 279
              | |||||  |  |   |  |    ||||||||    | | |   |  |  | |
Db        390 GTCGAACGCAAGTTCCTTTTCATACAGTGGGATGGAGAGGATGGCAGTAAAGTTGCTGTA 449

Qy        280 ATAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAA 339
              || | |   | |   | || ||   | |||     ||||||||  | ||||||||||||
Db        450 TTAAAGGAATTAGGAGAAGTAGACAAAGACAGACCTAATAACAGGATTAATGATGGCAAA 509
```

```
Qy        340 GCAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTC 391
              |||||||| | ||| |   | |   | || || |||| |  ||| | |  ||
Db        510 GCAGATCCTCGTGGGAGGTTGTTTGCTGGAACTATGGGTCATGAAGATCCTC 561


RESULT 3
BI611823
LOCUS       BI611823                670 bp    mRNA    linear   EST 07-SEP-2001
DEFINITION  RH18245.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH18245 5 similar to regucalcin:
            FBan0001803 GO:[] located on: X 11A3-11A4;: 08/17/2001, mRNA
            sequence.
ACCESSION   BI611823
VERSION     BI611823.1  GI:15507348
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 670)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003706: arm:3R [10470528,10725488]
            estimated-cyto:88D2-88E1: 08/17/2001
            Plate: RH.182  row: D  column: 9
            High quality sequence stop: 511.
FEATURES             Location/Qualifiers
     source          1. .670
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH18245"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      166 a    176 c    183 g    145 t
ORIGIN

  Query Match             7.7%;  Score 71.2;  DB 13;  Length 670;
  Best Local Similarity   47.9%;  Pred. No. 4.5e-09;
  Matches  273;  Conservative    0;  Mismatches  288;  Indels    9;  Gaps    2;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              | || || || || |||||||||  |||    ||||   | || | ||| ||| || | |
Db         93 TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 152

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAA------G 160
                  || ||   ||||        ||||| |||           ||||      |
Db        153 CCGCCGGCATTAATCGTTATGATTTCAAGCAGAACAAAGTGTACAGGGCTAAAATCGAGG 212

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
                || | | | |  || ||||| | ||  |||        ||       ||||  ||
Db        213 GCGAGATATTTGCATCGTTCATTCTGCCGGTTGAGAACAAACCGCAGGAGTTTGCCGTAG 272

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              | |  |  | |   |  | |||  |    |||||||| ||    || |   | |    ||
Db        273 GATGCGGTCTTCGTACGGTCATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAGGTCA 332

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
                |  |  | | ||  | ||  || | ||  | ||| | || ||||||| ||||
Db        333 CTCGCACCCTGTTCGAGGTGCAGCCGGACCTGAAGGAAAACCGCCTTAATGATGCCAAAA 392

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
              | ||||||  ||||  | | |      || || |||||    |||  |||   |    |
Db        393 CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGACATATTCA 449

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
                |    |  ||     | ||    |    ||| | |  | ||   ||      |   |
Db        450 CCCAATGGAAGGGTGAGCTCTACAGCTGGCAGGCCGGTGGACAGCCCAACGCTATCCGTA 509

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
              ||||  | |  ||| | ||||| || || |||   |    |    |||||  | || |  |
Db        510 GCAAGGTGGGCATATCCAATGGCCTGGCCTGGGATGTCAAGGCCAAGAAGTTCTACTTCA 569

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
              | ||  |    ||       |||   |  || || ||  ||      ||     || | |
Db        570 TCGACACCAACAACCACGAGGTATTGGCCTATGACTACCATCAGAGCACCGGCGCCGTAA 629

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAA 610
              |||| | |  | |||| |  |  |  | ||
Db        630 GCAACCCAAAGTCATTTTCGATCTGAGGAA 659


RESULT 4
BI617530
LOCUS       BI617530                635 bp    mRNA    linear   EST 07-SEP-2001
DEFINITION  RH47894.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH47894 5 similar to regucalcin:
            FBan0001803 GO:[] located on: X 11A3-11A4;: 08/18/2001, mRNA
            sequence.
ACCESSION   BI617530
VERSION     BI617530.1  GI:15513055
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003706: arm:3R [10470528,10725488]
            estimated-cyto:88D2-88E1: 08/18/2001
            Plate: RH.478  row: H  column: 10
            High quality sequence stop: 528.
FEATURES             Location/Qualifiers
     source          1. .635
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH47894"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
```

BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."
BASE COUNT 157 a 167 c 175 g 136 t
ORIGIN

Query Match 7.2%; Score 66.4; DB 13; Length 635;
Best Local Similarity 47.8%; Pred. No. 1e-07;
Matches 261; Conservative 0; Mismatches 276; Indels 9; Gaps 2;

```
Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
Db         93 TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 152

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAA------G 160
Db        153 CCGCCGGCATTAATCGTTATGATTTCAAGCAGAACAAAGTGTACAGGGCTAAAATCGAGG 212

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
Db        213 GCGAGATATTTGCATCGTTCATTCTGCCGGTTGAGAACAAACCGCAGGAGTTTGCCGTAG 272

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
Db        273 GATGCGGTCTTCGTACGGTCATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAGGTCA 332

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
Db        333 CTCGCACCCTGTTCGAGGTGCAGCCGGACCTGAAGGAAAACCGCCTTAATGATGCCAAAA 392

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
Db        393 CCGATCCCAATGGCCGTTTTACGGTGGCACCATGGC---CGACAGTGGCGACATATTCA 449

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
Db        450 CCCAATGGAAGGGTGAGCTCTACAGCTGGCAGGCCGGTGGACAGCCCAACGCTATCCGTA 509

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
Db        510 GCAAGGTGGGCATATCCAATGGCCTGGCCTGGGATGTCAAGGCCAAGAAGTTCTACTTCA 569

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
Db        570 TCGACACCAACAACCACGAGGTATTGGCCTATGACTACCATCAGAGCACCGGCGCCGTAA 629

Qy        581 GCAATC 586
Db        630 GCAACC 635
```

RESULT 5
AJ425659
LOCUS AJ425659 719 bp mRNA linear EST 31-JUL-2002
DEFINITION AJ425659 Salmo salar spleen adult Salmo salar cDNA clone s20C06, mRNA sequence.
ACCESSION AJ425659
VERSION AJ425659.1. GI:22033827
KEYWORDS EST.
SOURCE Atlantic salmon.
ORGANISM Salmo salar
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE 1 (bases 1 to 719)
AUTHORS Cairney,M. and Taggart,J.
TITLE Gene sequences in atlantic salmon (Salmo salar)
JOURNAL Unpublished (2002)
COMMENT Contact: Taggart JB
Institute of Aquaculture
University of Stirling
Airthrey Road, FK9 4LA, United Kingdom.
FEATURES Location/Qualifiers
source 1..719
/organism="Salmo salar"
/db_xref="taxon:8030"
/clone="s20C06"
/clone_lib="Salmo salar spleen adult"
/tissue_type="spleen"
/dev_stage="adult"
/note="Vector: Stratagene lambda Uni-Zap; country=United Kingdom"
BASE COUNT 192 a 168 c 197 g 161 t 1 others
ORIGIN

Query Match 7.1%; Score 65.4; DB 9; Length 719;
Best Local Similarity 47.3%; Pred. No. 2e-07;
Matches 301; Conservative 0; Mismatches 322; Indels 13; Gaps 3;

```
Qy        287 CTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATC 346
Db         28 CCATCGCTGTCATCGATGAAGACAAACCCAACAACCGTTTTAATGACGGCAAGGTGGACC 87

Qy        347 CCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGG 406
Db         88 CAGCTGGACGCCTCTTTGCAGGTACCATGGCGATGGAGGAACGTCCTACAGTTCTAGAGC 147

Qy        407 TCACTGGCAGTT---TATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCA 463
Db        148 TGAAGCAGGGTTCTCTGTTCTCCCTGAACAAGGACCACATTGTAGTCAAACATTTCAATC 207

Qy        464 ACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTG 523
Db        208 AGGTGGACATCTCAAACGGTCTGGACTGGTCTCTGGACCACAAAACCTTCTACTACATAG 267

Qy        524 ATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCA 583
Db        268 ACAGCCTCACCTACACTGTGGAAGCTTTCAACTATGACATCAACACCGGGAGTATCAGTA 327

Qy        584 ATCAACGGCCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAA 643
Db        328 ACCGGCGGATGGTGTATAAGATGGA-------GGAGGGGAGGGGATTCCTGATGGAATGT 380

Qy        644 CAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAA 703
Db        381 GTATTGACGCAGACGGTAGACTCTGGGTTGCCTGCTACAATGGAGGCAGAGTGATTCAGA 440

Qy        704 TCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCA 763
Db        441 T---TGACCCACAGACAGGGGTGCGTTTGCAGACAGTGAAGCTACCGGTGGATAAGACAA 497

Qy        764 CCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTC 823
Db        498 CATCATGTTGTTTCGGGGGTAGAGACTATTCAGATCTGTACGTGACCTCAGCCTGTAAGG 557

Qy        824 AGCTTGACGACAGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAG 883
Db        558 GAATGGACGAGGCAGACATGGCCAAGCAACCGCAGGCAGGCTGTGTCTTTAGGATCACAG 617

Qy        884 GTTTAGGCGTCAAAGGTTTCGCGGGAGTTAAAGTGA 919
Db        618 GCCTGGGTGCAAAGGGCATTCCTGCAAATTCATTCA 653
```

RESULT 6
BI617627
LOCUS BI617627 655 bp mRNA linear EST 07-SEP-2001
DEFINITION RH48022.5prime RH Drosophila melanogaster normalized Head pFlc-1 Drosophila melanogaster cDNA clone RH48022 5 similar to regucalcin: FBan0001803 GO:[] located on: X 11A3-11A4;: 08/18/2001, mRNA sequence.
ACCESSION BI617627
VERSION BI617627.1 GI:15513152
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

```
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003487: arm:X [11484037,11785087]
            estimated-cyto:10D4-11A4: 08/18/2001
            Plate: RH.480  row: B  column: 10
            High quality sequence stop: 538.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH48022"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      164 a    171 c    178 g    142 t
ORIGIN

  Query Match             7.1%;  Score 65.2;  DB 13;  Length 655;
  Best Local Similarity   47.6%;  Pred. No. 2.2e-07;
  Matches  261;  Conservative    0;  Mismatches  278;  Indels    9;  Gaps    2;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              | || || || || |||||||||  |||    ||||   | || | ||| |||  || |
Db         93 TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 152

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAA------G 160
                 || ||   ||||        |||||  |||           ||||      |
Db        153 CCGCCGGCATTAATCGTTATGATTTCAAGCAGAACAAAGTGTACAGGGCTAAAATCGAGG 212

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
                | | | | |  || |||||| | ||  |||       ||       ||||  ||
Db        213 GCGAGATATTTGCATCGTTCATTCTGCCGGTTGAGAACAAACCGCAGGAGTTTGCCGTAG 272

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              | |  |  | |   |  | |||  |    ||||||| ||    || |   | |    ||
Db        273 GATGCGGTCTTCGTACGGTCATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAGGTCA 332

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
                |  |  |  | ||  || | ||  | ||| | || ||||||| ||||
Db        333 CTCGCACCCTGTTCGAGGTGCAGCCGGACCTGAAGGAAAACCGCCTTAATGATGCCAAAA 392

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
              | ||||||  |||| | | |     || || |||||    |||  |||   |    |
Db        393 CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGACATATTCA 449

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
                |    | ||     | ||    |    ||| | |  | ||   ||      |   |
Db        450 CCCAATGGAAGGGTGAGCTCTACAGCTGGCAGGCCGGTGGACAGCCCAACGCTATCCGTA 509

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
              ||||  | |  ||| | ||||| || || ||   |    |    |||||  | || |  |
```

```
Db        510 GCAAGGTGGGCATATCCAATGGCCTGGCCTGAGATGTCAAGGCCAAGAAGTTCTACTTCA 569

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
              | ||  |   ||      |||   |  || || ||  ||      ||     || | |
Db        570 TCGACACCAACAACCACGAGGTATTGGCCTATGACTACAATCAGAGCACCGGCGCCGTAA 629

Qy        581 GCAATCAA 588
              |||| | |
Db        630 GCAACCCA 637


RESULT 7
C92366
LOCUS       C92366                   560 bp    mRNA    linear   EST 12-JUL-1999
DEFINITION  C92366 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSE652, mRNA sequence.
ACCESSION   C92366
VERSION     C92366.1  GI:3074242
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 560)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .560
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSE652"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT      195 a     83 c    104 g    178 t
ORIGIN

  Query Match             7.0%;  Score 65;  DB 14;  Length 560;
  Best Local Similarity   50.4%;  Pred. No. 2.4e-07;
  Matches  186;  Conservative    0;  Mismatches  180;  Indels    3;  Gaps    1;

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
              || |  |||||| | |  |||   | ||||   || ||  | || |||||||| |
Db         94 ATTGGTAATGGGTTGGTATGGTCAATTGATCAAAAAAAGTTTTACTATATTGATACACCA 153

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
              |     |   | || ||    |||||  | ||| |   |    |  |    ||||  ||
Db        154 ACTTTGAATATCGATGAAATGGATTACAACGCTGCCCAAGGCACTGTATCAAATCGTCGT 213

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
              |  |   |  |||||||||| | | |   |    || |  |||||||||   || ||||||
Db        214 CATTGTATATCTTTTGAAAACCGTAATTTC---GGCTTCCCAGATGGTATGACCATTGAT 270

Qy        652 GAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACC 711
                || |||||  | ||| ||||   ||   | ||     || |||  | |    || |
Db        271 AGTGAAGGTAAACTCTGGATTGCTCATTGGGAGGGTGGTAGAGTTACAAGATGGTGTCCA 330

Qy        712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTA 771
                      || | ||| |   |           |   |   |  |||||| |||
```

```
Db        331 ATCACTAAAAAACTTTTATTAACTGTAACTGTTCCAAATGTCTCAAGGGTCACTTCTTGT 390

Qy        772 GCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGAC 831
              || ||||| | | |  || |  |    ||  || | ||| ||||  | | | |   |||
Db        391 GCTTTTGGTGATTCTGATCTTCAATCTCTTTATATCACAACTGCAAGACGTGATGGTGAA 450

Qy        832 GACAGTTCT 840
                  ||||
Db        451 CCTGATTCT 459


RESULT 8
C92731
LOCUS       C92731                   606 bp    mRNA    linear   EST 12-JUL-1999
DEFINITION  C92731 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSF142, mRNA sequence.
ACCESSION   C92731
VERSION     C92731.1  GI:3074607
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 606)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .606
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSF142"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT      207 a     86 c    119 g    194 t
ORIGIN

  Query Match             7.0%;  Score 65;  DB 14;  Length 606;
  Best Local Similarity   50.4%;  Pred. No. 2.5e-07;
  Matches  186;  Conservative    0;  Mismatches  180;  Indels    3;  Gaps    1;

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
              || |  |||||| | |  |||   | ||||   || ||  | || |||||||||  |
Db        141 ATTGGTAATGGGTTGGTATGGTCAATTGATCAAAAAAAGTTTTACTATATTGATACACCA 200

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
              |      |   | || ||    |||||  | ||| |   |     |  |    ||||  ||
Db        201 ACTTTGAATATCGATGAAATGGATTACAACGCTGCCCAAGGCACTGTATCAAATCGTCGT 260

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
              |  |   | ||||||||||| | | |  |    || |  |||||||||   || ||||||
Db        261 CATTGTATATCTTTTGAAAACCGTAATTTC---GGCTTCCCAGATGGTATGACCATTGAT 317

Qy        652 GAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACC 711
              || |||||  | ||| ||||   ||   | ||     || |||  | |    ||  |
Db        318 AGTGAAGGTAAACTCTGGATTGCTCATTGGGAGGGTGGTAGAGTTACAAGATGGTGTCCA 377

Qy        712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTA 771
                  || | ||| |  .|          |   |   |   |||||| |||
```

```
Db        378 ATCACTAAAAAACTTTTATTAACTGTAACTGTTCCAAATGTCTCAAGGGTCACTTCTTGT 437

Qy        772 GCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGAC 831
              || ||||| | | |  || |  |    ||  || | ||| ||||  | | | |   |||
Db        438 GCTTTTGGTGATTCTGATCTTCAATCTCTTTATATCACAACTGCAAGACGTGATGGTGAA 497

Qy        832 GACAGTTCT 840
                  ||||
Db        498 CCTGATTCT 506


RESULT 9
C92540
LOCUS       C92540                   639 bp    mRNA    linear   EST 12-JUL-1999
DEFINITION  C92540 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSE589, mRNA sequence.
ACCESSION   C92540
VERSION     C92540.1  GI:3074416
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 639)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .639
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSE589"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT      233 a     87 c    120 g    199 t
ORIGIN

  Query Match             7.0%;  Score 65;  DB 14;  Length 639;
  Best Local Similarity   50.4%;  Pred. No. 2.5e-07;
  Matches  186;  Conservative    0;  Mismatches  180;  Indels    3;  Gaps    1;

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
              || |  |||||| | |  |||   | ||||   || ||  | || |||||||||  |
Db        150 ATTGGTAATGGGTTGGTATGGTCAATTGATCAAAAAAAGTTTTACTATATTGATACACCA 209

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
              |      |   | || ||    |||||  | ||| |   |     |  |    ||||  ||
Db        210 ACTTTGAATATCGATGAAATGGATTACAACGCTGCCCAAGGCACTGTATCAAATCGTCGT 269

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
              |  |   | ||||||||||| | | |  |    || |  |||||||||   || ||||||
Db        270 CATTGTATATCTTTTGAAAACCGTAATTTC---GGCTTCCCAGATGGTATGACCATTGAT 326

Qy        652 GAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACC 711
              || |||||  | ||| ||||   ||   | ||     || |||  | |    ||  |
Db        327 AGTGAAGGTAAACTCTGGATTGCTCATTGGGAGGGTGGTAGAGTTACAAGATGGTGTCCA 386

Qy        712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTA 771
                  || | ||| |  |          |   |   |   |||||| |||
```

```
Db        387 ATCACTAAAAAACTTTTATTAACTGTAACTGTTCCAAATGTCTCAAGGGTCACTTCTTGT 446

Qy        772 GCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGAC 831
              || ||||| | | |  || | |    ||  || | ||| |||| | | | |   |||
Db        447 GCTTTTGGTGATTCTGATCTTCAATCTCTTTATATCACAACTGCAAGACGTGATGGTGAA 506

Qy        832 GACAGTTCT 840
                  ||||
Db        507 CCTGATTCT 515


RESULT 10
C92417
LOCUS       C92417                   679 bp    mRNA    linear   EST 12-JUL-1999
DEFINITION  C92417 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSE888, mRNA sequence.
ACCESSION   C92417
VERSION     C92417.1  GI:3074293
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSE888"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT     255 a     94 c    124 g    206 t
ORIGIN

  Query Match             7.0%;  Score 65;  DB 14;  Length 679;
  Best Local Similarity   50.4%;  Pred. No. 2.6e-07;
  Matches  186;  Conservative    0;  Mismatches  180;  Indels    3;  Gaps    1;

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
              || |  |||||| | |  |||   | ||||   || ||  | || |||||||||| |
Db        193 ATTGGTAATGGGTTGGTATGGTCAATTGATCAAAAAAAGTTTTACTATATTGATACACCA 252

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
              |     |   | || ||    |||||  | ||| |    |  |    ||||  ||
Db        253 ACTTTGAATATCGATGAAATGGATTACAACGCTGCCCAAGGCACTGTATCAAATCGTCGT 312

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
              |  |   | ||||||||||| | | |    || |  |||||||||   || ||||||
Db        313 CATTGTATATCTTTTGAAAACCGTAATTTC---GGCTTCCCAGATGGTATGACCATTGAT 369

Qy        652 GAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACC 711
                || |||||  | ||| ||||   ||   | ||     || |||  | |    || |
Db        370 AGTGAAGGTAAACTCTGGATTGCTCATTGGGAGGGTGGTAGAGTTACAAGATGGTGTCCA 429

Qy        712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTA 771
                      || | ||| |   |           |   |   |   |||||| |||
```

```
Db        430 ATCACTAAAAAACTTTTATTAACTGTAACTGTTCCAAATGTCTCAAGGGTCACTTCTTGT 489

Qy        772 GCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGAC 831
              || ||||| | | |  || | |    ||  || | ||| |||| | | | |   |||
Db        490 GCTTTTGGTGATTCTGATCTTCAATCTCTTTATATCACAACTGCAAGACGTGATGGTGAA 549

Qy        832 GACAGTTCT 840
                  ||||
Db        550 CCTGATTCT 558


RESULT 11
C90258
LOCUS       C90258                   730 bp    mRNA    linear   EST 20-APR-1998
DEFINITION  C90258 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSI507, mRNA sequence.
ACCESSION   C90258
VERSION     C90258.1  GI:3059878
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     Developmental cDNA in Dictyostelium discoideum
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSI507"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT     263 a     101 c    135 g    231 t
ORIGIN

  Query Match             7.0%;  Score 65;  DB 14;  Length 730;
  Best Local Similarity   50.4%;  Pred. No. 2.6e-07;
  Matches  186;  Conservative    0;  Mismatches  180;  Indels    3;  Gaps    1;

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
              || |  |||||| | |  |||   | ||||   || ||  | || |||||||||| |
Db        253 ATTGGTAATGGGTTGGTATGGTCAATTGATCAAAAAAAGTTTTACTATATTGATACACCA 312

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
              |     |   | || ||    |||||  | ||| |    |  |    ||||  ||
Db        313 ACTTTGAATATCGATGAAATGGATTACAACGCTGCCCAAGGCACTGTATCAAATCGTCGT 372

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
              |  |   | ||||||||||| | | |    || |  |||||||||   || ||||||
Db        373 CATTGTATATCTTTTGAAAACCGTAATTTC---GGCTTCCCAGATGGTATGACCATTGAT 429

Qy        652 GAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACC 711
                || |||||  | ||| ||||   ||   | ||     || |||  | |    || |
Db        430 AGTGAAGGTAAACTCTGGATTGCTCATTGGGAGGGTGGTAGAGTTACAAGATGGTGTCCA 489

Qy        712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTA 771
                      || | ||| |   |           |   |   |   |||||| |||
Db        490 ATCACTAAAAAACTTTTATTAACTGTAACTGTTCCAAATGTCTCAAGGGTCACTTCTTGT 549

Qy        772 GCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGAC 831
              || ||||| | | |  || | |    ||  || | ||| |||| | | | |   |||
Db        550 GCTTTTGGTGATTCTGATCTTCAATCTCTTTATATCACAACTGCAAGACGTGATGGTGAA 609
```

```
Qy        832 GACAGTTCT 840
                  ||||
Db        610 CCTGATTCT 618


RESULT 12
BJ435579/c
LOCUS       BJ435579                743 bp    mRNA    linear   EST 13-MAR-2002
DEFINITION  BJ435579 Dictyostelium discoideum cDNA library, VF Dictyostelium
            discoideum cDNA clone ddv27g15 3', mRNA sequence.
ACCESSION   BJ435579
VERSION     BJ435579.1  GI:19410301
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the vegetative
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .743
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddv27g15"
                     /clone_lib="Dictyostelium discoideum cDNA library, VF"
                     /sex="mat A"
                     /dev_stage="Growth phase"
BASE COUNT      235 a    137 c    104 g    265 t      2 others
ORIGIN

  Query Match             7.0%;  Score 65;  DB 13;  Length 743;
  Best Local Similarity   50.4%;  Pred. No. 2.6e-07;
  Matches  186;  Conservative    0;  Mismatches  180;  Indels    3;  Gaps    1;

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
              || |  |||||| | |  |||   | ||||  || ||  | || |||||||| |
Db        470 ATTGGTAATGGGTTGGTATGGTCAATTGATCAAAAAAAGTTTTACTATATTGATACACCA 411

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
              |     |  | || ||    |||||  | ||| |   |    |  |    ||||  ||
Db        410 ACTTTGAATATCGATGAAATGGATTACAACGCTGCCCAAGGCACTGTATCAAATCGTCGT 351

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
              |  |   |  |||||||||| | | |   |    || |  |||||||||   || ||||||
Db        350 CATTGTATATCTTTTGAAAACCGTAATTTC---GGCTTCCCAGATGGTATGACCATTGAT 294

Qy        652 GAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACC 711
              || |||||  | ||| ||||   ||   | ||     || |||  | |    || |
Db        293 AGTGAAGGTAAACTCTGGATTGCTCATTGGGAGGGTGGTAGAGTTACAAGATGGTGTCCA 234

Qy        712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTA 771
                  || | ||| |   |           |   |   |   |||||| |||
Db        233 ATCACTAAAAAACTTTTATTAACTGTAACTGTTCCAAATGTCTCAAGGGTCACTTCTTGT 174

Qy        772 GCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGAC 831
              || ||||| | | |  || |  |    ||  || | ||| ||||  | | | |   |||
Db        173 GCTTTTGGTGATTCTGATCTTCAATCTCTTTATATCACAACTGCAAGACGTGATGGTGAA 114

Qy        832 GACAGTTCT 840
                  ||||
Db        113 CCTGATTCT 105
```

```
RESULT 13
AI663645
LOCUS       AI663645                555 bp    mRNA    linear   EST 10-MAY-1999
DEFINITION  ui99h04.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1890583 5' similar to SW:SM30_MOUSE Q64374 SENESCENCE MARKER
            PROTEIN-30 ;, mRNA sequence.
ACCESSION   AI663645
VERSION     AI663645.1  GI:4767228
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: ui99h04.x1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:974907
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 460.
FEATURES             Location/Qualifiers
     source          1. .555
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1890583"
                     /clone_lib="Sugano mouse liver mlia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT      158 a    106 c    146 g    143 t      2 others
ORIGIN

  Query Match             7.0%;  Score 64.4;  DB 9;  Length 555;
  Best Local Similarity   49.5%;  Pred. No. 3.6e-07;
  Matches  254;  Conservative    0;  Mismatches  247;  Indels   12;  Gaps    3;

Qy        306 ACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTATGGAC 365
              | | |  |||||||  || ||||||||||| || |  |||||   ||| |     |  |
Db          3 AGATAAGAAAAATAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGATACTTTGC 62

Qy        366 AGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGCAGTTTATA 422
                |||||  ||||||   ||  | |  |   | ||    |  | |  |   ||    || ||
Db         63 TGGTACCATGGCTGAGGAAACGGCCCCAGCTGTTCTTGAGCGGCACCAAGGGTCCTTGTA 122
```

```
Qy        423 TCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAATGG 482
Db        123 CTCCCTCTTTCCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATCTCCAATGG 182

Qy        483 GCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGAGT 542
Db        183 TTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCCTACACTGT 242

Qy        543 AGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTTAC 602
Db        243 GGATGCCTTTGACTATGACCTACAAACAGGACAGATTTCCAACCGCAGAATTGTTTACAA 302

Qy        603 TTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGTAA 662
Db        303 GATGGAAAAAGATGAACA------AATCCCAGATGGAATGTGCATTGATGCTGAGGGAAA 356

Qy        663 TTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCGGA 722
Db        357 GCTATGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTGGATCCTGAGACAGGGAA 416

Qy        723 AGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGCGG 782
Db        417 A---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGCTTTGGAGG 473

Qy        783 TCCGAATTTGGATGAACTGCATGTAACATCTGC 815
Db        474 GAAAGATTACTCTNGAATGTATGTGACCTGTGC 506


RESULT 14
AV890554/c
LOCUS       AV890554                730 bp    mRNA    linear   EST 09-NOV-2001
DEFINITION  AV890554 Nori Satoh unpublished cDNA library, cleavage stage embryo
            Ciona intestinalis cDNA clone rcicl28p05 3', mRNA sequence.
ACCESSION   AV890554
VERSION     AV890554.1  GI:16879635
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="rcicl28p05"
                     /clone_lib="Nori Satoh unpublished cDNA library, cleavage
                     stage embryo"
                     /tissue_type="whole animal"
                     /dev_stage="cleavage stage embryo"
                     /note="Vector: pBluescript SK"
BASE COUNT      197 a    166 c    134 g    233 t
ORIGIN

  Query Match             6.9%;  Score 63.4;  DB 10;  Length 730;
  Best Local Similarity   52.6%;  Pred. No. 7.5e-07;
  Matches  214;  Conservative    0;  Mismatches  181;  Indels   12;  Gaps    3;

Qy        412 GGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCT 471
Db        718 GGCAAGCTTTATTGTCTTGACATTGATCATTCCGTAAAGACAAAGGTTTATCCGGTAGAT 659

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
Db        658 ATTTCCAATGGTCTTTCATGGACGA------GCAATACAATGTACTACTGCGACTCGTTG 605

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
Db        604 AAGCTTACGATAGACGCTTACGATTACGATGTAACCACGGGCGAAATAAAAAATATGAGG 545

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
Db        544 GAAATTGTGAAGTTTG---ATCGTGAGAAGGAAGGAGTTCCAGATGGCCACTGTATTGAC 488

Qy        652 GAGGAGGGTAATTTATGGGTTGC---CGTTTTCCAAGGACAGCGAATTATTAAAATCAGT 708
Db        487 ACCGATGGTAACTTATGGGTGGCAATGTTTTTCACCGGCCAAGTAATCAAAATTGACCCG 428

Qy        709 ACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCT 768
Db        427 AGGACAGGAGAGAAGCTGCAATATGTGAAAGTAAGCGACATCGCTCTCAAAACAACGTCG 368

Qy        769 GTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGC 815
Db        367 GTATGTTTCGGAGGTCCGAACTTGGACGAAATGTACGTCACGTCGGC 321


RESULT 15
BJ371993/c
LOCUS       BJ371993                513 bp    mRNA    linear   EST 08-MAR-2002
DEFINITION  BJ371993 Dictyostelium discoideum cDNA library, CF Dictyostelium
            discoideum cDNA clone ddc10c20 3', mRNA sequence.
ACCESSION   BJ371993
VERSION     BJ371993.1  GI:19281376
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the culmination
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .513
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddc10c20"
                     /clone_lib="Dictyostelium discoideum cDNA library, CF"
                     /sex="mat A"
                     /dev_stage="Culmination stage"
BASE COUNT      166 a    103 c     82 g    162 t
ORIGIN

  Query Match             6.8%;  Score 62.4;  DB 13;  Length 513;
  Best Local Similarity   50.9%;  Pred. No. 1.3e-06;
  Matches  175;  Conservative    0;  Mismatches  166;  Indels    3;  Gaps    1;

Qy        472 ATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGG 531
Db        341 ATTGGTAATGGGTTGGTATGGTCAATTGATCAAAAAAAGTTTTACTATATTGATACACCA 282

Qy        532 AAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGG 591
Db        281 ACTTTGAATATCGATGAAATGGATTACAACGCTGCCCAAGGCACTGTATCAAATCGTCGT 222

Qy        592 CCATTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGAT 651
```

```
Db        221 CATTGTATATCTTTTGAAAACCGTAATTTC---GGCTTCCCAGATGGTATGACCATTGAT 165

Qy        652 GAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACC 711

Db        164 AGTGAAGGTAAACTCTGGATTGCTCATTGGGAGGGTGGTAGAGTTACAAGATGGTGTCCA 105

Qy        712 CAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTA 771

Db        104 ATCACTAAAAAACTTTTATTAACTGTAACTGTTCCAAATGTCTCAAGGGTCACTTCTTGT 45

Qy        772 GCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGC 815

Db         44 GCTTTTGGTGATTCTGATCTTCAATCTCTTTATATCACAACTGC 1
```

Search completed: August 1, 2003, 13:20:58
Job time : 1776 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:01:21 ; Search time 292 Seconds
(without alignments)
7126.186 Million cell updates/sec

Title: US-10-089-986-1
Perfect score: 924
Sequence: 1 atggggccagttgttgaaaa.....cgggagttaaagtgaagcta 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	863	93.4	865	22	AAF86444	Firefly luciferin
	2	325.2	35.2	924	24	ABL49712	Luciola lateralis
	3	320.4	34.7	930	24	ABL41998	Nucleotide sequenc
	4	98.6	10.7	912	23	ABL12833	Drosophila melanog
	5	90.4	9.8	1216	16	AAQ87295	Rat regucalcin cDN
	6	80.8	8.7	1573	24	ABL99968	Mouse SMP30 encodi
	7	74.4	8.1	1356	16	AAQ90035	Human SMP30 gene.
	8	74.4	8.1	1356	20	AAX40196	Senescence marker
	9	74.4	8.1	1356	24	ABN95150	Gene #1648 used to
c	10	70.4	7.6	2970	23	ABL12832	Drosophila melanog
	11	67	7.3	1086	23	ABL03467	Drosophila melanog
	12	67	7.3	1094	23	ABL19751	Drosophila melanog
	13	67	7.3	1119	23	ABL03469	Drosophila melanog
	14	52.8	5.7	3554	23	ABL03468	Drosophila melanog
	15	52.8	5.7	4624	23	ABL03466	Drosophila melanog
	16	52.8	5.7	6625	23	ABL19750	Drosophila melanog
	17	46.8	5.1	1346	22	AAH57498	Human liver cell s
	18	42.2	4.6	4065	19	AAV42941	Nucleic acid encod
	19	42.2	4.6	4065	20	AAX39676	Renal cancer assoc
	20	42.2	4.6	4065	24	ABL59542	Human ROCK1 cDNA S
	21	42.2	4.6	4739	18	AAT78203	Physiologically ac
c	22	41.2	4.5	2365589	24	ABA90521	Genomic sequence o
	23	40.6	4.4	2641	23	ABL28434	Drosophila melanog
	24	39.8	4.3	428	24	ABQ58878	Human colon cancer
	25	39.8	4.3	4590	22	AAH24065	Yeast AOD9604-asso
	26	38.8	4.2	684707	24	ABQ67196	Listeria innocua c
c	27	38.8	4.2	3011208	24	ABQ69245	Listeria innocua D
c	28	38.6	4.2	11812	22	AAS45501	Chemically pretrea
c	29	38.6	4.2	11812	22	AAS46741	Tumour suppressor
c	30	38.6	4.2	11812	24	ABL34118	Human immune syste
c	31	38.6	4.2	11812	24	ABK28431	DNA transcription
	32	38	4.1	66986	22	AAF28542	Genomic fragment #
c	33	37.8	4.1	1993	24	ABQ70703	Listeria monocytog
c	34	36.6	4.0	3078	24	ABN59645	Novel human coding
	35	35.8	3.9	1114	22	AAF91977	B thuringiensis 44
	36	35.8	3.9	1152	18	AAV06394	B. thuringiensis 4
	37	35.8	3.9	1152	21	AAA96766	DNA sequence encod
	38	35.8	3.9	1241	22	AAF91952	B thuringiensis 44
	39	35.8	3.9	1521	22	AAF92018	B thuringiensis 58
	40	35.8	3.9	2132	18	AAV06392	B. thuringiensis 1
	41	35.8	3.9	2132	21	AAA96764	Nucleotide sequenc
	42	35.8	3.9	2132	22	AAF91950	B thuringiensis 14
	43	35.2	3.8	32038	22	AAS30028	Human lung antigen
	44	35.2	3.8	32038	22	AAK89980	Human digestive sy
	45	35.2	3.8	32038	22	AAK90877	Human digestive sy

ALIGNMENTS

RESULT 1
AAF86444
ID AAF86444 standard; DNA; 865 BP.
XX
AC AAF86444;
XX
DT 26-JUN-2001 (first entry)
XX
DE Firefly luciferin regenerating protein coding sequence.
XX
KW Firefly; luciferin regenerating protein; oxyluciferin; luciferase; ds.
XX
OS Photinus pyralis.
XX
FH Key Location/Qualifiers
FT CDS 2..865
FT /*tag= a
FT /product= "Luciferin regenerating protein"
FT /transl_except= (pos:782..784, aa:Leu)
FT /note= "No start or stop codon given"
XX
PN WO200125426-A1.
XX
PD 12-APR-2001.
XX
PF 22-SEP-2000; 2000WO-JP06527.
XX
PR 06-OCT-1999; 99JP-0285258.
XX
PA (KIKK) KIKKOMAN CORP.
XX
PI Kurosawa K, Kajiyama N;

```
XX
DR   WPI; 2001-266307/27.
DR   P-PSDB; AAB82087.
XX
PT   Luciferin regenerating protein and gene encoding it useful for
PT   regenerating expensive luciferin from oxyluciferin and D-cysteine -
XX
PS   Disclosure; Page 14-15; 21pp; Japanese.
XX
CC   The present sequence is the coding sequence from firefly for a protein
CC   that regenerates luciferin by using oxyluciferin and D-cysteine. The
CC   encoded protein can be used for regeneration of luciferin, a substrate
CC   for luciferase, used for ATP (adenosine triphosphate) assays in both
CC   medical and food hygiene areas.
XX
SQ   Sequence 865 BP; 267 A; 163 C; 195 G; 239 T; 1 other;

  Query Match             93.4%;  Score 863;  DB 22;  Length 865;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-242;
  Matches  864;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         61 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 61

Qy        121 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 121

Qy        181 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 181

Qy        241 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTC 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTC 241

Qy        301 GAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTA 301

Qy        361 TGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTA 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTA 361

Qy        421 TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAAT 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAAT 421

Qy        481 GGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGA 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGA 481

Qy        541 GTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTT 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTT 541

Qy        601 ACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGT 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGT 601

Qy        661 AATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCG 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCG 661

Qy        721 GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGC 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGC 721

Qy        781 GGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGACAGTTCT 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGACAGTTCT 781

Qy        841 TTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTCAAAGGT 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTCAAAGGT 841

Qy        901 TTCGCGGGAGTTAAAGTGAAGCTA 924
               ||||||||||||||||||||||||
Db        842 TTCGCGGGAGTTAAAGTGAAGCTA 865


RESULT 2
ABL49712
ID   ABL49712 standard; DNA; 924 BP.
XX
AC   ABL49712;
XX
DT   29-MAY-2002  (first entry)
XX
DE   Luciola lateralis luciferin regenerating protein encoding DNA SEQ:1.
XX
KW   Luciola lateralis; luciferin; regeneration; Japanese firefly;
KW   adenosine triphosphate; ATP; medical science; food hygiene;
KW   luciferase; luminescence; gene; ds.
XX
OS   Luciola lateralis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..924
FT                   /*tag= a
FT                   /product= "Luciferin regeneration protein"
XX
PN   WO200210384-A1.
XX
PD   07-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-JP06455.
XX
PR   28-JUL-2000; 2000JP-0228227.
XX
PA   (KIKK ) KIKKOMAN CORP.
XX
PI   Hirokawa K,  Kurosawa K,  Kajiyama N;
XX
DR   WPI; 2002-227086/28.
DR   P-PSDB; ABB06339.
XX
PT   Luciola lateralis-originated genes encoding proteins capable of
PT   regenerating luciferin especially from oxyluciferin, useful for
PT   producing recombinant DNAs and transformants to give proteins useful in
PT   assaying adenosine triphosphate  -
XX
PS   Claim 9; Page 17-18; 23pp; Japanese.
XX
CC   The present sequence encodes a protein capable of regenerating
CC   luciferin, which is isolated from Luciola lateralis (Japanese firefly).
CC   The gene encoding the protein capable of regenerating luciferin can be
CC   used for producing recombinant DNAs and transformants, which can be
CC   used for the production of proteins useful in assaying adenosine
CC   triphosphate (ATP) in medical sciences and food hygiene. The protein
CC   can be added to the luciferin-luciferase reaction system to sustain
CC   luminescence and reduce their amounts used.
XX
SQ   Sequence 924 BP; 325 A; 146 C; 178 G; 275 T; 0 other;

  Query Match             35.2%;  Score 325.2;  DB 24;  Length 924;
  Best Local Similarity   60.2%;  Pred. No. 8.2e-85;
  Matches  556;  Conservative    0;  Mismatches  364;  Indels    3;  Gaps    1;

Qy          1 ATGGGGCCAGTTGTTGAAAAAATTGCAGAACTTGGCAAGTATACGGTTGGAGAAGGTCCT 60
              |||  |||||||  |||||  | ||  | |||  | |  |||  |       | |||||  |||||
Db          1 ATGTCGCCAGTTATTGAACAGATCACTGAAGTAGACAATTTCCAAATCGGAGAGGGTCCA 60

Qy         61 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 120
               |||||||||   |||||  ||  |  ||  |||||  ||  ||  |   ||||  |||  |  |  ||
```

```
Db        61 CACTGGGATACCGAAACACAAAGTTTGTATTTTGTGGATATTCTAGAAAAATCCATACAC 120

Qy       121 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 180
             ||||| ||||| ||    ||| || | ||        |   | ||||||     |||||
Db       121 AAATACGTACCATCGACAAAACAACATACTAAAATGATTTTGAATAAACGTCCGTCTTTT 180

Qy       181 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 240
             ||||| ||  |    | | |  |||  || |||||  | ||||| ||||| || ||   |
Db       181 ATTATACCAATAAAAGAAACATCTGATCGGTTTGTCATAAGTTTAGAACGAGACATTTGC 240

Qy       241 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTC 300
              ||||||| |||||||| |||||||||  | || |      ||||| | || ||||  ||
Db       241 GTTCTTACTTGGGATGGTGTTAGTGCCACGCCAAGTCATTTAGAAACAATAGTTACCGTT 300

Qy       301 GAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTA 360
             ||  |    ||  ||   || | | ||||||||||| |||||||| |  |||||||| |
Db       301 GATACGGGAATCGAAGGAAATACATTCAATGATGGTAAAGCAGATGCGTTTGGCAATTTG 360

Qy       361 TGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTA 420
             ||| |||||||| |   || | |   | | |      |  |      | |    |   |
Db       361 TGGGCAGGTACATTATATAGTAAATTCGATATTGAAAAACAAGGTCCTAATACAGGAACA 420

Qy       421 TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAAT 480
             |   |      | || |||| || ||| ||  ||       || ||     || |||||
Db       421 TTGTACAGCCTGTCTAATAAGCAGTTAAGAAAACATATTTCTAATATCTTCCTATCAAAT 480

Qy       481 GGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGA 540
             || || || |||| ||| || |  || ||| | |||| ||| || ||    ||||||| |
Db       481 GGCCTAGCCTGGAATAAAGACTCAAAAAAATTTTATTTTATCGACTCCAATAAAAGAACA 540

Qy       541 GTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTT 600
              ||||  |||  |||||||||| ||   | |||   ||    |||   |  ||||| |||
Db       541 ATAGATCAGTTTGATTATGATTCTGAAAATTTAATTATATCAAATTGTCAACCATTGTTT 600

Qy       601 ACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGT 660
             ||| | || || ||||   | |  || |  || ||||  ||||| || |||||  | | |
Db       601 ACTCTGGACAAACATGGGATACAGGGTTTACCTGATGCCCAAACGATAGATGAAAATGAT 660

Qy       661 AATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCG 720
             ||||||||||||||   |  | | ||||     | |||| || ||  ||||  | |||||
Db       661 AATTTATGGGTTGCTATAGTTCGAGGAGGAAAAGTTATAAATATTGGTACTAAGCAACCC 720

Qy       721 GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGC 780
             |||   || || | |    | || || || ||   |  | | || || ||    |||||
Db       721 GAATCTTTGCTTGGTGTTATTAACATGCCTGAAAGTTTGATAACTTCAGTTTGTTTTGGA 780

Qy       781 GGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGACAGTTCT 840
             ||  | || ||||| |||||  |||||||  || ||||| |  |     | || |   ||
Db       781 GGATCAAAGTTGGACGAACTTTATGTAACGACTTCTGGTATCAAAGAGTATGAAA---CT 837

Qy       841 TTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTCAAAGGT 900
                | ||   |||| || ||     | |||||||| || || ||||| || ||||||
Db       838 GACTCTACAAAACTAGTAAAGGGTGGATTGTACAGAGTTACTGGATTAGGTGTTAAAGGT 897

Qy       901 TTCGCGGGAGTTAAAGTGAAGCT 923
             ||  | | |   | | | |  ||
Db       898 TTACCCGCACACAGATTCAGTCT 920


RESULT 3
ABL41998
ID   ABL41998 standard; DNA; 930 BP.
XX
AC   ABL41998;
XX
DT   11-JUN-2002  (first entry)
XX
DE   Nucleotide sequence of a protein capable of regenerating luciferin.
XX
KW   Luciferin regenerating protein; luciferen; oxyluciferin; luminescence;
KW   ATP assay; gene; ss.
XX
OS   Luciola cruciata.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..930
FT                   /*tag=  a
FT                   /product= "Luciferin regenerating protein"
XX
PN   WO200210383-A1.
XX
PD   07-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-JP06454.
XX
PR   28-JUL-2000; 2000JP-0228226.
XX
PA   (KIKK ) KIKKOMAN CORP.
XX
PI   Hirokawa K,  Kurosawa K,  Kajiyama N;
XX
DR   WPI; 2002-180080/23.
DR   P-PSDB; ABB09720.
XX
PT   New Luciola cruciata-originated genes encoding proteins capable of
PT   regenerating luciferin especially from oxyluciferin, for producing
PT   recombinant DNAs and transformants -
XX
PS   Claim 9; Page 17-18; 23pp; Japanese.
XX
CC   The present sequence encodes a protein capable of regenerating
CC   luciferin. The protein can especially regenerate luciferen from
CC   oxyluciferin. The protein can be added to the luciferin-luciferase
CC   reaction system to sustain luminescence and reduce the amounts used.
CC   The polynucleotide sequence is useful for producing recombinant
CC   transformants, which are useful for the production of proteins useful
CC   in assaying ATP in medical sciences and food hygiene.
XX
SQ   Sequence 930 BP; 305 A; 150 C; 203 G; 272 T; 0 other;

  Query Match             34.7%;  Score 320.4;  DB 24;  Length 930;
  Best Local Similarity   60.6%;  Pred. No. 2.1e-83;
  Matches  563;  Conservative    0;  Mismatches  357;  Indels    9;  Gaps    2;

Qy         1 ATGGGGCCAGTTGTTGAAAAAATTGCAGAACTTGGCAAGTATACGGTTGGAGAAGGTCCT 60
             ||||  |||  ||||||| |||| |  ||| | |||| |||  | ||| |||| ||||
Db         1 ATGGCTCCAACTGTTGAACAAATAGTGGAATTGGGCACTTATTTGCTTGCAGAAAGTCCA 60

Qy        61 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 120
             || |||||  | |||||||| |  || || || || || |  ||||  | | || |  |
Db        61 CATTGGGACGACGAAACTCAAAGTTTGTACTTTGTGGATATTGTAGGAAGATCTGTAAAC 120

Qy       121 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 180
             || || || ||  ||   ||||   ||||       ||| | ||||||      || ||
Db       121 AAGTACGTGCCGACTACCAAAACGCACACACAACTGAAATTTGATAAAAATCCATCGTTT 180

Qy       181 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 240
             || |||||  |    ||||  | |  |||||| | || ||||| ||||| |||||
Db       181 ATCATTCCTGTCAAAGGATGTTCCGATCGTTTTATTGTGAGTTTAGAACGAGAAATTAAT 240

Qy       241 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAA------GCTATTGTT 294
             ||||||||||||||||  | |||| |||| || |  |  ||||||      ||| |  ||
Db       241 CTTCTTACATGGGATGGTGCTAGTTCTGCTCCAAGTAAGATAGAAAAAATCGCTGTATTT 300

Qy       295 AATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGC 354
             ||       ||  | |  |    | || ||| | ||||| || |||||||| || |||||
Db       301 GATAATACTCCTGAAAAAAGTGAAAATAGATTAAATGACGGTAAAGCAGACCCTCTTGGA 360

Qy       355 AATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGC 414
             |||||  ||| | || || |||  ||| |    | | ||    |   ||||  |   | | |
Db       361 AATCTGTGGGCTGGAACGATGAATATGGGTTCAGATCATACGACAGGAACACCCGTTCGT 420

Qy       415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
             |    |            || ||||  | || |||   ||||  |    |  |   ||||
Db       421 GGCACGTTGTCGAGTTTATCTAATAAGCAAGTGAAAGAACACGTGTCTGAAGTTTGTATA 480
```

```
Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
Db        481 TCCAATGGTCTTGCTTGGAGTAAAGATTTAAAAAAGTTTTATTACATTGATTCTGCTGTT 540

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
Db        541 AGACAAGTCGATCAATTTGATTTTGATGCAAAAAATTTATCACTTTCTAACCGACAACCG 600

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
Db        601 TTATTTACATTTGACAAACATGGAATTATGGGTTCACCTGATGGACAGACTATAGATGCC 660

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
Db        661 GAAGGAAACTTATGGGTAGCCACATGTCAAGGTGATAAAGTTTTAAAAATTGATACTAGT 720

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
Db        721 ACTCCGGAAACCTTACTTGGAATTGTCGAGATTCCAGAACATCAGGTGACATCAGTTTGC 780

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
Db        781 ATTGGTGGAGCGGAATTAAATGTATTGTATGTTACTACTGCCAGTATAAAACTTCCCG-- 838

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
Db        839 -GTGCTGATGAAACAAAACCAATGAAGGGTGCTATCTATAAAGTAACTGGACTTGGTGTG 897

Qy        895 AAAGGTTTCGCGGGAGTTAAAGTGAAGCT 923
Db        898 AAAGGATTGCCAGGAGATCGAGTTAAGTT 926
```

RESULT 4
ABL12833
ID ABL12833 standard; cDNA; 912 BP.
XX
AC ABL12833;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32981.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75:
DR P-PSDB; ABB68730.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 32981; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 912 BP; 231 A; 246 C; 238 G; 197 T; 0 other;

Query Match 10.7%; Score 98.6; DB 23; Length 912;
Best Local Similarity 49.2%; Pred. No. 1.7e-18;
Matches 384; Conservative 0; Mismatches 379; Indels 18; Gaps 4;

```
Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
Db         44 TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 103

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAA------G 160
Db        104 CCGCCGGCATTAATCGTTATGATTTCAAGCAGAACAAAGTGTACAGGGCTAAAATCGAGG 163

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
Db        164 GCGAGATATTTGCATCGTTCATTCTGCCGGTTGAGAACAAACCGCAGGAGTTTGCCGTAG 223

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
Db        224 GATGCGGTCTTCGTACGGTCATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAGGTCA 283

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
Db        284 CTCGCACCCTGTTCGAGGTGCAGCCGGACCTGAAGGAAAACCGCCTTAATGATGCCAAAA 343

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
Db        344 CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGACATATTCA 400

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
Db        401 CCCAATGGAAGGGTGAGCTCTACAGCTGGCAGGCCGGTGGACAGCCCAACGCTATCCGTA 460

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
Db        461 GCAAGGTGGGCATATCCAATGGCCTGGCCTGGGATGTCAAGGCCAAGAAGTTCTACTTCA 520

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
Db        521 TCGACACCAACAACCACGAGGTATTGGCCTATGACTACAATCAGAGCACCGGCGCCGTAA 580

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAA------GCATGAAGTGCCTGGATATCCAG 634
Db        581 GCAACCCAAAGGTCATCTTCGATCTGAGGAAGATTCGGCCCGAAGGACCATTGTTCCCTG 640

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
Db        641 ATGGCATGACCGTAGACACCGATGGCAATATCTACGTGGCCACCTTCAATGGTGGCACCG 700

Qy        695 TTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATC 754
Db        701 TCTTCAAGGTCAATCCAAGCACCGGTAAAATTCTGCTGGAGA---TCAAAATTCCAACCA 757

Qy        755 CTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTG 814
Db        758 CCCAAATCACCTCGGTGGCTTTTGGAGGTCCCAATTTGGATATTTTGTATGTGACAACCG 817

Qy        815 C 815
Db        818 C 818
```

RESULT 5

```
AAQ87295
ID   AAQ87295 standard; cDNA; 1216 BP.
XX
AC   AAQ87295;
XX
DT   31-JAN-1996  (first entry)
XX
DE   Rat regucalcin cDNA.
XX
KW   Calcium binding protein; regucalcin; rat; ds.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             80..979
FT                   /*tag=  a
FT                   /product= regucalcin
FT                   /transl_except= pos:362..364, aa:Glu
FT                   /transl_except= pos:500..502, aa:His
XX
PN   JP07123985-A.
XX
PD   16-MAY-1995.
XX
PF   09-NOV-1993;   93JP-0279349.
XX
PR   09-NOV-1993;   93JP-0279349.
XX
PA   (DAII-) DAIICHI KAKAGU YAKUHIN KK.
PA   (YAMA/) YAMAGUCHI M.
XX
DR   WPI; 1995-211628/28.
DR   P-PSDB; AAR75416.
XX
PT   DNA encoding regucalcin - useful for the recombinant production of
PT   regucalcin, a calcium binding protein
XX
PS   Claim 1; Page 5-6; 6pp; Japanese.
XX
CC   A rat liver cDNA clone coding for a calcium binding protein of mol.
CC   wt. 33388 was isolated and sequenced (AAQ87295). The protein was
CC   designated "regucalcin".
XX
SQ   Sequence 1216 BP; 338 A; 231 C; 309 G; 338 T; 0 other;

  Query Match             9.8%;  Score 90.4;  DB 16;  Length 1216;
  Best Local Similarity   50.2%;  Pred. No. 4.9e-16;
  Matches  307;  Conservative    0;  Mismatches  292;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || || | | | |  ||||| ||  || ||||||||||| || |  |||||   ||| |
Db        356 GTGGATCAAGATAAGAAAAACAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 415

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414
                |   | ||||| ||||||   ||  | |  |   | ||    |  | | |  |   ||
Db        416 TACTTTGCTGGTACCATGGCTGAGGAAACCGCCCCAGCTGTTCTGGAGCGGCACCAAGGG 475

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
              || ||     |    ||||| | |  || || |   ||   | | |  | | |||
Db        476 TCCTTGTACTCCCTTTTTCCTGATGACAGTGTGAAGAAATACTTTAACCAAGTGGATATC 535

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
               | |||||  | |  |||      ||    || |   | || || |||||      |
Db        536 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 595

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
                | || || |  |  || |||||     | |||  |   ||   ||| |   || |
Db        596 TACACTGTGGATGCCTTTGACTATGACCTGCCAACAGGACAGATTTCCAACCGCAGGACT 655

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
              | |  |   | |||||  |||||        |   ||||||||       |||||||
Db        656 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGTT 709

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
              ||||| ||  | |||||| |||  || | | |||    || | |||    |   | |  |
Db        710 GAGGGGAAGCTTTGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTAGATCCTGAG 769

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
                |  | ||     |||| | || || ||  | || | |  | |    || ||
Db        770 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 826

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
              ||||| ||   | |||    |||| || | || |||| |||  |   |  | |   || |
Db        827 TTTGGAGGGAAGGATTACTCTGAAATGTACGTGACATGTGCCAGGGATGGGATGAGCGCC 886

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
                | |    |        | |  |||  || | | ||   |||||||| | || |||
Db        887 GAAGGTCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTTGGGGTC 946

Qy        895 AAAGGTTTCGC 905
              |||||  | ||
Db        947 AAAGGAATTGC 957


RESULT 6
ABL99968
ID   ABL99968 standard; cDNA; 1573 BP.
XX
AC   ABL99968;
XX
DT   14-AUG-2002  (first entry)
XX
DE   Mouse SMP30 encoding cDNA SEQ ID NO 2.
XX
KW   Mouse; senescence marker protein 30; cancer; aging; senescence; SMP30;
KW   cytostatic; gene; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             73..972
FT                   /*tag=  a
FT                   /product= "SMP30"
XX
PN   WO200234041-A1.
XX
PD   02-MAY-2002.
XX
PF   22-OCT-2001; 2001WO-JP09243.
XX
PR   23-OCT-2000; 2000JP-0322234.
XX
PA   (MARU/) MARUYAMA N.
PA   (KASA/) KASAHARA Y.
XX
PI   Maruyama N,  Kasahara Y;
XX
DR   WPI; 2002-405445/43.
DR   P-PSDB; ABB83787.
XX
PT   Non-human animals with defective senescence marker protein 30 function,
PT   useful for producing antibody, histopathological diagnosis and onset
PT   mechanism of disease due to aging, and in screening side-effects during
PT   drug development  -
XX
PS   Disclosure; Page 38-40; 44pp; Japanese.
XX
CC   The invention relates to a senescence marker protein 30 (SMP30) non-human
CC   animal having a defect in the SMP30 function. The animals are useful for
CC   producing antibody, evaluating its biological function, histopathological
CC   diagnosis and onset mechanism of disease due to aging, and in screening
CC   side-effects during development of drug and cosmetics including for
CC   cancer treatment.
XX
SQ   Sequence 1573 BP; 453 A; 288 C; 366 G; 466 T; 0 other;
```

```
  Query Match             8.7%;  Score 80.8;  DB 24;  Length 1573;
  Best Local Similarity   49.3%;  Pred. No. 3.6e-13;
  Matches  301;  Conservative    0;  Mismatches  298;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||   | | |  |||||||| || ||||||||||| || |  |||||   ||| |
Db        349 GTGGATGAAGATAAGAAAAATAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 408

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414
                | | ||||| ||||||   ||  | |  |  | ||    |  | |  |   ||
Db        409 TACTTTGCTGGTACCATGGCTGAGGAAACGGCCCCAGCTGTTCTTGAGCGGCACCAAGGG 468

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
                || ||     |     ||||| | |  || || |   ||     | |  | | |||
Db        469 TCCTTGTACTCCCTCTTTCCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATC 528

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
              | |||||  | |  |||      ||   || |   | || || |||||      |
Db        529 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 588

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
                | || || | |  || |||||       |||  |   ||   ||| |   |
Db        589 TACACTGTGGATGCCTTTGACTATGACCTACAAACAGGACAGATTTCCAACCGCAGAATT 648

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
              | |  |   | |||||  |||||         |  ||||||||       |||||||
Db        649 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGCT 702

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
              ||||| ||  |||||||| |||  || | | |||    || | |||    |  | |  |
Db        703 GAGGGAAAGCTATGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTGGATCCTGAG 762

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
                | | ||     |||| | || || ||  | || | |  ( |     || ||
Db        763 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 819

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
              ||||| ||    |||   |||| || |||| || | |||  |   |  | | |  |
Db        820 TTTGGAGGGAAAGATTACTCTGAAATGTATGTGACCTGTGCCAGGGATGGGTTGAATGCT 879

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
                  |    |       | |  |||  || | | ||   |||||||| | || |||
Db        880 GAAGGCCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTCGGAGTC 939

Qy        895 AAAGGTTTCGC 905
              |||||  | ||
Db        940 AAAGGAATTGC 950
```

```
RESULT 7
AAQ90035
ID   AAQ90035 standard; cDNA to mRNA; 1356 BP.
XX
AC   AAQ90035;
XX
DT   03-JAN-1996  (first entry)
XX
DE   Human SMP30 gene.
XX
KW   SMP30; marker protein; ageing; organ development; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             94..993
FT                   /*tag= a
FT                   /product= SMP30 marker protein
XX
PN   JP07097399-A.
XX
PD   11-APR-1995.
XX
PF   29-SEP-1993;   93JP-0265681.
XX
PR   29-SEP-1993;   93JP-0265681.
XX
PA   (FJRE ) FUJI REBIO KK.
XX
DR   WPI; 1995-175363/23.
DR   P-PSDB; AAR74219.
XX
PT   Novel polypeptide for detecting human ageing marker protein SMP30 -
PT   for monitoring liver and kidney development in new-born babies
XX
PS   Claim 3; Page 6-7; 10pp; Japanese.
XX
CC   AAQ90035 is the SMP30 gene. It encodes the human ageing marker protein,
CC   SMP30 (AAR74219). Human SMP30 is found in human organs, tissues, blood,
CC   urine and cerebrospinal fluid. The blood concentration of SMP30 is
CC   known to increase with renal and hepatic deficiencies and to decrease
CC   with age. It is therefore useful in the monitoring of renal or hepatic
CC   deficiencies and for the monitoring of the development of the liver
CC   and kidneys in newborn babies.
XX
SQ   Sequence 1356 BP; 373 A; 274 C; 343 G; 366 T; 0 other;
```

```
  Query Match             8.1%;  Score 74.4;  DB 16;  Length 1356;
  Best Local Similarity   48.6%;  Pred. No. 2.5e-11;
  Matches  297;  Conservative    0;  Mismatches  302;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||     |||  ||||| ||  |  ||||||||||| || |  ||||||   || |
Db        370 GTGGATAACGACAAGAAAAACAATCGCTTCAATGATGGGAAGGTGGATCCCGCCGGGAGG 429

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCG---TAGGACCGGTCACTGGC 414
                  |   | || || ||||||   ||  | | ||   | |   | |  | |   ||
Db        430 TACTTTGCTGGCACCATGGCTGAGGAAACAGCTCCAGCAGTTCTTGAGCGGCACCAGGGG 489

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
                  | ||     |     ||||| |  | || |||| | ||     | |  | |  ||
Db        490 GCCCTGTACTCCCTCTTTCCTGATCACCACGTGAAAAAGTACTTTGACCAGGTGGACATT 549

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
                | |||||  | |  |||      ||   || |   | ||||| |||||      |
Db        550 TCCAATGGTTTGGATTGGTCGCTAGACCACAAAATCTTCTATTACATTGACAGCCTGTCC 609

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
                  || || |  |  || |||||       |||  |  |||  ||| |   |
Db        610 TACTCCGTGGATGCCTTTGACTATGACCTGCAGACAGGACAGATCTCCAACCGCAGAAGT 669

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
                | |  |   | |||||| | |||       |   ||||||||       |||||||
Db        670 GTTTACAAGCTAGAAAAGGAAGAA------CAAATCCCAGATGGAATGTGTATTGATGCT 723

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
              |||||  ||  | |||||  |||  || | | |||    || | |||     |  | |
Db        724 GAGGGGAAGCTCTGGGTGGCCTGTTACAATGGAGGAAGAGTGATTCGTTTAGATCCTGTG 783

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
                | | ||     |||  | || || ||  | || | |  |   || ||
Db        784 ACAGGGAAA---AGACTTCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 840

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
              |||||  ||   |||||    |||| || |||| || |  ||  |   |  | | ||| |
Db        841 TTTGGAGGGAAGAATTACTCTGAAATGTATGTGACCTGCGCCCGGGATGGGATGGACCCC 900

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
                  | |  |     |     |||     | | ||   |||| ||| | || |||
Db        901 GAGGGTCTTTTGAGGCAACCTGAAGCTGGTGGAATTTTCAAGATAACTGGTCTGGGGGTC 960

Qy        895 AAAGGTTTCGC 905
              |||||  | ||
Db        961 AAAGGAATTGC 971
```

RESULT 8

```
AAX40196
ID   AAX40196 standard; DNA; 1356 BP.
XX
AC   AAX40196;
XX
DT   02-JUL-1999  (first entry)
XX
DE   Senescence marker protein SMP30 gene.
XX
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9904265-A2.
XX
PD   28-JAN-1999.
XX
PF   15-JUL-1998;   98WO-US14679.
XX
PR   22-JUN-1998;   98US-0102322.
PR   17-JUL-1997;   97US-0896164.
PR   10-OCT-1997;   97US-0061599.
PR   10-OCT-1997;   97US-0061765.
PR   10-OCT-1997;   97US-0948705.
PR   11-OCT-1997;   97GB-0021697.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Chen Y,  Gout I,  Gure A,  O'Hare M,  Obata Y,  Old LJ;
PI   Pfreundschuh M,  Sahin U,  Scanlan MJ,  Stockert E;
PI   Tureci O;
XX
DR   WPI; 1999-132448/11.
DR   P-PSDB; AAY06995.
XX
PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
XX
PS   Claim 67; Page 775; 787pp; English.
XX
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
XX
SQ   Sequence 1356 BP; 373 A; 274 C; 343 G; 366 T; 0 other;

  Query Match             8.1%;  Score 74.4;  DB 20;  Length 1356;
  Best Local Similarity   48.6%;  Pred. No. 2.5e-11;
  Matches  297;  Conservative    0;  Mismatches  302;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||     |||  ||||| ||  |  ||||||||||| || |  ||||||   || |
Db        370 GTGGATAACGACAAGAAAAACAATCGCTTCAATGATGGGAAGGTGGATCCCGCCGGGAGG 429

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCG---TAGGACCGGTCACTGGC 414
                |  | || || ||||||   ||  | | ||  | |   | |  | |  |  ||
Db        430 TACTTTGCTGGCACCATGGCTGAGGAAACAGCTCCAGCAGTTCTTGAGCGGCACCAGGGG 489

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
                 | ||     |     ||||| |  | || |||| | ||     | |  | |  ||
Db        490 GCCCTGTACTCCCTCTTTCCTGATCACCACGTGAAAAAGTACTTTGACCAGGTGGACATT 549

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
               | |||||  | |  |||       ||    || |   | |||||  |||||       |
Db        550 TCCAATGGTTTGGATTGGTCGCTAGACCACAAAATCTTCTATTACATTGACAGCCTGTCC 609

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
                  || || |  |  || |||||        |||  |   |||  ||| |   |
Db        610 TACTCCGTGGATGCCTTTGACTATGACCTGCAGACAGGACAGATCTCCAACCGCAGAAGT 669

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
               | |  |   | ||||||  | |||          |   ||||||||       |||||||
Db        670 GTTTACAAGCTAGAAAAGGAAGAA------CAAATCCCAGATGGAATGTGTATTGATGCT 723

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
              |||||  ||  | |||||  |||  || | |  |||    || | |||    |   | |
Db        724 GAGGGGAAGCTCTGGGTGGCCTGTTACAATGGAGGAAGAGTGATTCGTTTAGATCCTGTG 783

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
                |  | ||     |||  | || || ||  | || | |  | |    || ||
Db        784 ACAGGGAAA---AGACTTCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 840

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
              |||||  ||   |||||    |||| || |||| || |  ||  |   |  | | |||  |
Db        841 TTTGGAGGGAAGAATTACTCTGAAATGTATGTGACCTGCGCCCCGGGATGGGATGGACCCC 900

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
                | |    |         |    |||     | | ||   |||| ||| | || |||
Db        901 GAGGGTCTTTTGAGGCAACCTGAAGCTGGTGGAATTTTCAAGATAACTGGTCTGGGGGTC 960

Qy        895 AAAGGTTTCGC 905
              |||||  | ||
Db        961 AAAGGAATTGC 971


RESULT 9
ABN95150
ID   ABN95150 standard; DNA; 1356 BP.
XX
AC   ABN95150;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #1648 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30589.
XX
PR   02-OCT-2000; 2000US-237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer,
PT   hepatocellular carcinoma or metastatic liver tumor in a patient,
PT   involves detecting the level of expression of two or more genes in a
PT   liver tissue sample  -
XX
PS   Claim 1; SEQ ID NO 1648; 298pp; English.
XX
```

```
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1356 BP; 373 A; 274 C; 343 G; 366 T; 0 other;

  Query Match             8.1%;  Score 74.4;  DB 24;  Length 1356;
  Best Local Similarity   48.6%;  Pred. No. 2.5e-11;
  Matches  297;  Conservative    0;  Mismatches  302;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||     |||  ||||| ||  |  |||||||||| || |  ||||||   || |
Db        370 GTGGATAACGACAAGAAAAACAATCGCTTCAATGATGGGAAGGTGGATCCCGCCGGGAGG 429

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCG---TAGGACCGGTCACTGGC 414
                |   | || || ||||||   ||  | | ||   | |   | |  | |  |   ||
Db        430 TACTTTGCTGGCACCATGGCTGAGGAAACAGCTCCAGCAGTTCTTGAGCGGCACCAGGGG 489

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
                | ||     |      ||||| |  | || |||| | ||     | |  | |  ||
Db        490 GCCCTGTACTCCCTCTTTCCTGATCACCACGTGAAAAAGTACTTTGACCAGGTGGACATT 549

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
                | |||||  | |  |||     ||   || |   | ||||| |||||       |
Db        550 TCCAATGGTTTGGATTGGTCGCTAGACCACAAAATCTTCTATTACATTGACAGCCTGTCC 609

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
                 || || |  |  || |||||       |||  |   |||  ||| |  |
Db        610 TACTCCGTGGATGCCTTTGACTATGACCTGCAGACAGGACAGATCTCCAACCGCAGAAGT 669

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
              | |  |   | ||||||  | |||         |   ||||||||       |||||||
Db        670 GTTTACAAGCTAGAAAAGGAAGAA------CAAATCCCAGATGGAATGTGTATTGATGCT 723

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
              |||||  ||  | |||||  |||  || | | |||    || | |||   .  |  | |
Db        724 GAGGGGAAGCTCTGGGTGGCCTGTTACAATGGAGGAAGAGTGATTCGTTTAGATCCTGTG 783

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
                |  | ||    |||  | || || ||  | || | |  |  |  ||  ||
Db        784 ACAGGGAAA---AGACTTCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 840

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
              ||||| ||   |||||    |||| || |||| || |  ||  |   |  | | |||  |
Db        841 TTTGGAGGGAAGAATTACTCTGAAATGTATGTGACCTGCGCCCGGGATGGGATGGACCCC 900

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
                | |    |        |    |||     | | ||   |||| ||| | || |||
Db        901 GAGGGTCTTTTGAGGCAACCTGAAGCTGGTGGAATTTTCAAGATAACTGGTCTGGGGGTC 960

Qy        895 AAAGGTTTCGC 905
              |||||  | ||
Db        961 AAAGGAATTGC 971


RESULT 10
ABL12832/c
ID   ABL12832 standard; cDNA; 2970 BP.
XX
AC   ABL12832;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 32978.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB68729.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 32978; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2970 BP; 893 A; 558 C; 574 G; 945 T; 0 other;

  Query Match             7.6%;  Score 70.4;  DB 23;  Length 2970;
  Best Local Similarity   47.3%;  Pred. No. 5.3e-10;
  Matches  320;  Conservative    0;  Mismatches  341;  Indels   15;  Gaps    3;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              | || || || || |||||||||  |||    ||||   | || | ||| |||  || |
Db       1927 TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 1868

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAA------G 160
                  || ||   ||||        |||| |||             ||||        |
Db       1867 CCGCCGGCATTAATCGTTATGATTTCAAGCAGAACAAAGTGTACAGGGCTAAAATCGAGG 1808

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220.
              || | | |  || |||||| | ||  |||         ||       ||||  ||
Db       1807 GCGAGATATTTGCATCGTTCATTCTGCCGGTTGAGAACAAACCGCAGGAGTTTGCCGTAG 1748

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              | |  | | |   | | |||  |    |||||||| ||    || |   | |    ||
Db       1747 GATGCGGTCTTCGTACGGTCATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAGGTCA 1688

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
                  |  |  | ||  | ||· || | ||  | ||| | || ||||||| ||||
Db       1687 CTCGCACCCTGTTCGAGGTGCAGCCGGACCTGAAGGAAAACCGCCTTAATGATGCCAAAA 1628

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
              | ||||||  ||||  | | |     || || |||||    |||  |||   |    |
Db       1627 CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGACATATTCA 1571
```

```
Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
              |     |  ||     | ||    |    ||| | |  | ||   ||      |   |
Db       1570 CCCAATGGAAGGGTGAGCTCTACAGCTGGCAGGCCGGTGGACAGCCCAACGCTATCCGTA 1511

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
              ||||  | |  ||| | ||||| || || |||  |    |    |||||  | || |  |
Db       1510 GCAAGGTGGGCATATCCAATGGCCTGGCCTGGGATGTCAAGGCCAAGAAGTTCTACTTCA 1451

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
              | ||  |    ||       |||   |  || || ||  ||      ||     || | |
Db       1450 TCGACACCAACAACCACGAGGTATTGGCCTATGACTACAATCAGAGCACCGGCGCCGTAA 1391

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAA------GCATGAAGTGCCTGGATATCCAG 634
              |||| | |  |    | ||   | |    ||      ||  ||||  ||    |  || |
Db       1390 GCAACCCAAAGGTCATCTTCGATCTGAGGAAGATTCGGCCCGAAGGACCATTGTTCCCTG 1331

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
              ||||    ||  | ||    || || ||| | |  || |||   ||| | ||
Db       1330 ATGGCATGACCGTAGACACCGATGGCAATATCTACGTGGCCACCTTCAATGGTGGCACCG 1271

Qy        695 TTATTAAAATCAGTAC 710
              |  | ||  ||| | |
Db       1270 TCTTCAAGGTCAATCC 1255
```

RESULT 11
ABL03467
ID ABL03467 standard; cDNA; 1086 BP.
XX
AC ABL03467;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 4883.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB59364.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 4883; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1086 BP; 245 A; 308 C; 296 G; 237 T; 0 other;

Query Match 7.3%; Score 67; DB 23; Length 1086;
Best Local Similarity 46.6%; Pred. No. 3.3e-09;
Matches 365; Conservative 0; Mismatches 400; Indels 18; Gaps 4;

```
Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              | || || ||||| || ||||||   |  |  |||| | | || | ||| ||    |  |
Db        131 TGGGCGAGGGTCCCCATTGGGATGTGGCCAGGCAGAGCCTGTACTACGTGGATTTGGAGG 190

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAA------AATACACGTTTTGTAAAG 160
                  |  | |      || |    | | ||||| |       |  | ||         | |
Db        191 CAGGCAGCCTGCTCCGCTACGACTATGCGCAGAACAAGGTCTACAAGACAAAGATCGAGG 250

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
               || |   |||      ||| |  | ||  | |  |||   ||       || |  |||
Db        251 GCGAAACCTTGGCCGGATTCGTGCTGCCGGTGGAGGGACGTCCGCAGGAATTCGCCGTCG 310

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              | |  |  ||     | |  ||  | |  ||||||||||||     |     | |
Db        311 GCTGCGGTCGACGCGTGGTGATCGTCAACTGGGATGGCGTCTCGCCCAGCGCCAAGGTGG 370

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
              |     |  |  |  | ||  | ||||  ||  | || ||  |  | || ||  |  |
Db        371 TGCGCACACTGTTCGAGGTGCAGCCACTGATGGAGAAGAATCGTTTGAACGACGCCAAGG 430

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
               |||||||  |||     | |      || || |||         |  | |     ||
Db        431 TTGATCCCCGTGGTCGCTTCTTTGGCGGCACCATGCGCTACATTGGCGATGAGTTCGAGT 490

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
              |    |   |||     | || |  |  | ||| | |    ||||     ||  | ||
Db        491 TCCGTCAC---GGCGAGCTGTACCGCTGGGAGGCCGGTGGCCAGGTGTCGGTGATCAAGG 547

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
              || |  | |  ||  | ||||| || || |||     |  |    |||||  | || || |
Db        548 GCGATGTGGGCATCTCCAATGGACTGGCATGGGACGAGAAGGCCAAGAAGTTCTACTACA 607

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
              | ||| |     |       ||  |   ||| || ||||||      ||          |
Db        608 TCGATACCACCGACTACGAGGTGAAGTCGTATGACTATGATTTCGAGACCGGCGTGGCTA 667

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAAGCAT------GAAGTGCCTGGATATCCAG 634
              ||||||    |    |||| | | |    ||| ||        | |  | |      || |
Db        668 GCAATCCCAAGGTTATATTCAATCTGCGCAAGAATAGTCCCAAGGATCATCTGCTGCCCG 727

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
              |||| |  || || |||   ||||| ||  | |  || ||| |||  | ||      |
Db        728 ATGGCCTGACCATCGATACCGAGGGCAACCTGTATGTGGCCACCTTCAATGGCGCCACCA 787

Qy        695 TTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATC 754
              |    ||  | | | || |         |  |  | || || | |   ||  | |||
Db        788 TCTACAAGGTTAATCCCAACACTGGCAAGATTCTGCTTGAGATC---AAGTTCCCAACCA 844

Qy        755 CTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTG 814
                 ||| | ||||| |  || || || || || || || |||||    ||| | || ||  |||
Db        845 AACAGATTACCTCCGCCGCCTTCGGTGGCCCCAACTTGGACATCCTGTACGTGACCACTG 904

Qy        815 CTG 817
              | |
Db        905 CCG 907
```

RESULT 12
ABL19751
ID ABL19751 standard; DNA; 1094 BP.
XX
AC ABL19751;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 10726.

XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 10726; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1094 BP; 245 A; 293 C; 292 G; 264 T; 0 other;

Query Match 7.3%; Score 67; DB 23; Length 1094;
Best Local Similarity 46.6%; Pred. No. 3.3e-09;
Matches 365; Conservative 0; Mismatches 400; Indels 18; Gaps 4;

```
Qy        47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
             | || || ||||| || ||||||   |  |  |||| | | || | ||| ||    |  |
Db       139 TGGGCGAGGGTCCCCATTGGGATGTGGCCAGGCAGAGCCTGTACTACGTGGATTTGGAGG 198

Qy       107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAA------AATACACGTTTTGTAAAG 160
                 |   | |     || |    | | |||||  |      |  | ||        | |
Db       199 CAGGCAGCCTGCTCCGCTACGACTATGCGCAGAACAAGGTCTACAAGACAAAGATCGAGG 258

Qy       161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
               || |   |||     ||| |  | ||  | |  |||   ||        || |  |||
Db       259 GCGAAACCTTGGCCGGATTCGTGCTGCCGGTGGAGGGACGTCCGCAGGAATTCGCCGTCG 318

Qy       221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
             | |  |  ||     | |  ||  | |  ||||||||||      |     | |
Db       319 GCTGCGGTCGACGCGTGGTGATCGTCAACTGGGATGGCGTCTCGCCCAGCGCCAAGGTGG 378

Qy       281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
             |     |  |  |  | ||  | ||||  ||  | || ||  |  | || || | ||| |
Db       379 TGCGCACACTGTTCGAGGTGCAGCCACTGATGGAGAAGAATCGTTTGAACGACGCCAAGG 438

Qy       341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
             |||||||  |||     | |     || || |||          |  | |     ||
Db       439 TTGATCCCCGTGGTCGCTTCTTTGGCGGCACCATGCGCTACATTGGCGATGAGTTCGAGT 498

Qy       401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
               |   |  |||   | || |  |  | ||| | |    ||||    ||  | ||
Db       499 TCCGTCAC---GGCGAGCTGTACCGCTGGGAGGCCGGTGGCCAGGTGTCGGTGATCAAGG 555

Qy       461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
             || |  | |  ||  | ||||| || || |||    |  |    |||||  | || || |
Db       556 GCGATGTGGGCATCTCCAATGGACTGGCATGGGACGAGAAGGCCAAGAAGTTCTACTACA 615

Qy       521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
             | ||| |     |       ||  |   ||| || ||||||      ||          |
Db       616 TCGATACCACCGACTACGAGGTGAAGTCGTATGACTATGATTTCGAGACCGGCGTGGCTA 675

Qy       581 GCAATCAACGGCCATTATTTACTTTTGAAAAGCAT------GAAGTGCCTGGATATCCAG 634
             ||||||    |    |||| | | |     ||| ||       | |  | |      || |
Db       676 GCAATCCCAAGGTTATATTCAATCTGCGCAAGAATAGTCCCAAGGATCATCTGCTGCCCG 735

Qy       635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
             |||| |  || || |||   ||||| ||  | |  || |||   ||| | ||        |
Db       736 ATGGCCTGACCATCGATACCGAGGGCAACCTGTATGTGGCCACCTTCAATGGCGCCACCA 795

Qy       695 TTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATC 754
             |    ||  | | | || |        |  |  | || || | |   ||  | |||
Db       796 TCTACAAGGTTAATCCCAACACTGGCAAGATTCTGCTTGAGATC---AAGTTCCCAACCA 852

Qy       755 CTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTG 814
               ||| | ||||| |  || || || || || || ||||    ||| | || ||  |||
Db       853 AACAGATTACCTCCGCCGCCTTCGGTGGCCCCAACTTGGACATCCTGTACGTGACCACTG 912

Qy       815 CTG 817
             | |
Db       913 CCG 915
```

RESULT 13
ABL03469
ID ABL03469 standard; cDNA; 1119 BP.
XX
AC ABL03469;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 4889.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB59366.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 4889; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins

```
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1119 BP; 266 A; 302 C; 310 G; 241 T; 0 other;

  Query Match             7.3%;  Score 67;  DB 23;  Length 1119;
  Best Local Similarity   46.6%;  Pred. No. 3.3e-09;
  Matches  365;  Conservative    0;  Mismatches  400;  Indels   18;  Gaps    4;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
Db        164 TGGGCGAGGGTCCCCATTGGGATGTGGCCAGGCAGAGCCTGTACTACGTGGATTTGGAGG 223

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAA------AATACACGTTTTGTAAAG 160
Db        224 CAGGCAGCCTGCTCCGCTACGACTATGCGCAGAACAAGGTCTACAAGACAAAGATCGAGG 283

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
Db        284 GCGAAACCTTGGCCGGATTCGTGCTGCCGGTGGAGGGACGTCCGCAGGAATTCGCCGTCG 343

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
Db        344 GCTGCGGTCGACGCGTGGTGATCGTCAACTGGGATGGCGTCTCGCCCAGCGCCAAGGTGG 403

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
Db        404 TGCGCACACTGTTCGAGGTGCAGCCACTGATGGAGAAGAATCGTTTGAACGACGCCAAGG 463

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
Db        464 TTGATCCCCGTGGTCGCTTCTTTGGCGGCACCATGCGCTACATTGGCGATGAGTTCGAGT 523

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
Db        524 TCCGTCAC---GGCGAGCTGTACCGCTGGGAGGCCGGTGGCCAGGTGTCGGTGATCAAGG 580

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
Db        581 GCGATGTGGGCATCTCCAATGGACTGGCATGGGACGAGAAGGCCAAGAAGTTCTACTACA 640

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
Db        641 TCGATACCACCGACTACGAGGTGAAGTCGTATGACTATGATTTCGAGACCGGCGTGGCTA 700

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAAGCAT------GAAGTGCCTGGATATCCAG 634
Db        701 GCAATCCCAAGGTTATATTCAATCTGCGCAAGAATAGTCCCAAGGATCATCTGCTGCCCG 760

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
Db        761 ATGGCCTGACCATCGATACCGAGGGCAACCTGTATGTGGCCACCTTCAATGGCGCCACCA 820

Qy        695 TTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATC 754
Db        821 TCTACAAGGTTAATCCCAACACTGGCAAGATTCTGCTTGAGATC---AAGTTCCCAACCA 877

Qy        755 CTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTG 814
Db        878 AACAGATTACCTCCGCCGCCTTCGGTGGCCCCAACTTGGACATCCTGTACGTGACCACTG 937

Qy        815 CTG 817
Db        938 CCG 940


RESULT 14
ABL03468
ID   ABL03468 standard; cDNA; 3554 BP.
XX
AC   ABL03468;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 4886.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB59365.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 4886; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3554 BP; 964 A; 784 C; 781 G; 1025 T; 0 other;

  Query Match             5.7%;  Score 52.8;  DB 23;  Length 3554;
  Best Local Similarity   46.0%;  Pred. No. 8.2e-05;
  Matches  300;  Conservative    0;  Mismatches  337;  Indels   15;  Gaps    3;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
Db       1542 TGGGCGAGGGTCCCCATTGGGATGTGGCCAGGCAGAGCCTGTACTACGTGGATTTGGAGG 1601

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAA------AATACACGTTTTGTAAAG 160
Db       1602 CAGGCAGCCTGCTCCGCTACGACTATGCGCAGAACAAGGTCTACAAGACAAAGATCGAGG 1661

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
Db       1662 GCGAAACCTTGGCCGGATTCGTGCTGCCGGTGGAGGGACGTCCGCAGGAATTCGCCGTCG 1721

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
Db       1722 GCTGCGGTCGACGCGTGGTGATCGTCAACTGGGATGGCGTCTCGCCCAGCGCCAAGGTGG 1781

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
Db       1782 TGCGCACACTGTTCGAGGTGCAGCCACTGATGGAGAAGAATCGTTTGAACGACGCCAAGG 1841

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
Db       1842 TTGATCCCCGTGGTCGCTTCTTTGGCGGCACCATGCGCTACATTGGCGATGAGTTCGAGT 1901
```

```
Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
              |    |   |||    | || |  |  | ||| | |    ||||     ||  | ||
Db       1902 TCCGTCAC---GGCGAGCTGTACCGCTGGGAGGCCGGTGGCCAGGTGTCGGTGATCAAGG 1958

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
              || |  | |  ||  | ||||| || || |||    |  |    |||||  | || || |
Db       1959 GCGATGTGGGCATCTCCAATGGACTGGCATGGGACGAGAAGGCCAAGAAGTTCTACTACA 2018

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
              | ||| |     |       ||  |   ||| || ||||||     ||          |
Db       2019 TCGATACCACCGACTACGAGGTGAAGTCGTATGACTATGATTTCGAGACCGGCGTGGCTA 2078

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAAGCAT------GAAGTGCCTGGATATCCAG 634
              ||||||    |    |||| | | |    ||| ||        | |  | |       || |
Db       2079 GCAATCCCAAGGTTATATTCAATCTGCGCAAGAATAGTCCCAAGGATCATCTGCTGCCCG 2138

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |||| |  || || |||   ||||| ||  | |  || |||   ||| | ||
Db       2139 ATGGCCTGACCATCGATACCGAGGGCAACCTGTATGTGGCCACCTTCAATGG 2190
```

RESULT 15
ABL03466
ID ABL03466 standard; cDNA; 4624 BP.
XX
AC ABL03466;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 4880.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB59363.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 4880; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4624 BP; 1242 A; 1013 C; 1017 G; 1352 T; 0 other;

Query Match 5.7%; Score 52.8; DB 23; Length 4624;
Best Local Similarity 46.0%; Pred. No. 9.3e-05;
Matches 300; Conservative 0; Mismatches 337; Indels 15; Gaps 3;

```
Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              | || || ||||| || ||||||    |  |  |||| | | || | ||| ||    |  |
Db       2612 TGGGCGAGGGTCCCCATTGGGATGTGGCCAGGCAGAGCCTGTACTACGTGGATTTGGAGG 2671

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAA------AATACACGTTTTGTAAAG 160
                 |   | |      || |    | | |||||  |      |  | ||       | |
Db       2672 CAGGCAGCCTGCTCCGCTACGACTATGCGCAGAACAAGGTCTACAAGACAAAGATCGAGG 2731

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
              || |   |||      ||| |  | ||  | |  |||   ||       || |  |||
Db       2732 GCGAAACCTTGGCCGGATTCGTGCTGCCGGTGGAGGGACGTCCGCAGGAATTCGCCGTCG 2791

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              | |  |  ||      | |  ||  | |  ||||||||||||     |     | |
Db       2792 GCTGCGGTCGACGCGTGGTGATCGTCAACTGGGATGGCGTCTCGCCCAGCGCCAAGGTGG 2851

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
              |     |  |  |  | ||  | ||||  ||  |  || ||  |  | || || | ||| |
Db       2852 TGCGCACACTGTTCGAGGTGCAGCCACTGATGGAGAAGAATCGTTTGAACGACGCCAAGG 2911

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
              |||||||  |||      | |      || || |||          |  | |     ||
Db       2912 TTGATCCCCGTGGTCGCTTCTTTGGCGGCACCATGCGCTACATTGGCGATGAGTTCGAGT 2971

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
              |    |   |||    | || |  |  | ||| | |    ||||     ||  | ||
Db       2972 TCCGTCAC---GGCGAGCTGTACCGCTGGGAGGCCGGTGGCCAGGTGTCGGTGATCAAGG 3028

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
              || |  | |  ||  | ||||| || || |||    |  |    |||||  | || || |
Db       3029 GCGATGTGGGCATCTCCAATGGACTGGCATGGGACGAGAAGGCCAAGAAGTTCTACTACA 3088

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
              | ||| |     |       ||  |   ||| || ||||||     ||          |
Db       3089 TCGATACCACCGACTACGAGGTGAAGTCGTATGACTATGATTTCGAGACCGGCGTGGCTA 3148

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAAGCAT------GAAGTGCCTGGATATCCAG 634
              ||||||    |    |||| | | |    ||| ||        | |  | |       || |
Db       3149 GCAATCCCAAGGTTATATTCAATCTGCGCAAGAATAGTCCCAAGGATCATCTGCTGCCCG 3208

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGG 686
              |||| |  || || |||   ||||| ||  | |  || |||   ||| | ||
Db       3209 ATGGCCTGACCATCGATACCGAGGGCAACCTGTATGTGGCCACCTTCAATGG 3260
```

Search completed: August 1, 2003, 12:06:43
Job time : 305 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 12:01:54 ; Search time 2636 Seconds
(without alignments)
10201.439 Million cell updates/sec

Title: US-10-089-986-1
Perfect score: 924
Sequence: 1 atggggccagttgttgaaaa.....cgggagttaaagtgaagcta 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		923	99.9	927	3	AB062786	AB062786 Photinus
2		325.2	35.2	924	3	AB072447	AB072447 Luciola l
3		320.4	34.7	930	3	AB072448	AB072448 Luciola c
4		98.6	10.7	1040	3	AY071130	AY071130 Drosophil
5		97.2	10.5	1150	3	AY028616	AY028616 Calliphor
6		93.8	10.2	1022	3	AB029490	AB029490 Drosophil
7		91	9.8	992	3	AF326959	AF326959 Calliphor
8		90.4	9.8	1216	6	E09237	E09237 cDNA insert
9		89.8	9.7	1155	3	AB036903	AB036903 Sarcophag
10		88.8	9.6	1594	10	RATCBP1	D38467 Rat mRNA fo
11		85.6	9.3	1594	10	RNSMP30A	X69021 R.norvegicu
12		85.6	9.3	1605	10	AB037934	AB037934 Rattus no
13		80.8	8.7	1109	10	D86217	D86217 Mus musculu
14		80.8	8.7	1573	10	MMU28937	U28937 Mus musculu
15		80.8	8.7	1598	10	BC012710	BC012710 Mus muscu
16		77.4	8.4	1126	5	AB037936	AB037936 Xenopus l
17		74.4	8.1	1356	6	AX409001	AX409001 Sequence
18		74.4	8.1	1356	6	E09013	E09013 cDNA encodi
19		74.4	8.1	1356	9	HUMSMP30	D31815 Human mRNA
20		74.4	8.1	1375	9	AB028125	AB028125 Homo sapi
21		74.4	8.1	1438	9	AB032064	AB032064 Homo sapi
22		73.6	8.0	1088	5	AB033368	AB033368 Xenopus l
23		73.2	7.9	269050	1	AP000987	AP000987 Sulfolobu
24		71.6	7.7	1394	4	AB035446	AB035446 Bos tauru
25		70.4	7.6	49817	2	AC013939	AC013939 Drosophil
26	c	70.4	7.6	167195	3	AC007808	AC007808 Drosophil
27	c	70.4	7.6	254961	3	AE003706	AE003706 Drosophil
28		69.4	7.5	888	1	SAU05664	U05664 Sulfolobus
29		67.8	7.3	1357	4	AB035445	AB035445 Oryctolag
30		67	7.3	1108	3	AY118643	AY118643 Drosophil
31		67	7.3	1208	3	AB036904	AB036904 Drosophil
32		65	7.0	56152	2	AC116963	AC116963 Dictyoste
33		62.8	6.8	633	3	AB079885	AB079885 Sarcophag
34		62.8	6.8	10105	1	AE006864	AE006864 Sulfolobu
35	c	59.8	6.5	286756	1	AP000990	AP000990 Sulfolobu
36		59.2	6.4	1221	5	AB037935	AB037935 Gallus ga
37	c	52.8	5.7	76808	2	AC014187	AC014187 Drosophil
38	c	52.8	5.7	186714	3	AC023688	AC023688 Drosophil
39		52.8	5.7	301051	3	AE003487	AE003487 Drosophil
40		50.4	5.5	12835	1	AE006895	AE006895 Sulfolobu
41		46.6	5.0	24650	2	AC023731	AC023731 Drosophil
42		45.6	4.9	1141	6	AX083744	AX083744 Sequence
43		44.4	4.8	100886	2	AP003831	AP003831 Oryza sat
44	c	44.4	4.8	164237	2	AP005199	AP005199 Oryza sat
45	c	43.4	4.7	7218	6	I66494	I66494 Sequence 14

ALIGNMENTS

RESULT 1
AB062786
LOCUS AB062786 927 bp mRNA linear INV 27-SEP-2001
DEFINITION Photinus pyralis mRNA for luciferin regenerating enzyme, complete cds.
ACCESSION AB062786
VERSION AB062786.1 GI:14331151
KEYWORDS
SOURCE Photinus pyralis cDNA to mRNA.
ORGANISM Photinus pyralis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia; Cantharoidea; Lampyridae; Photinus.
REFERENCE 1
AUTHORS Gomi,K. and Kajiyama,N.

```
  TITLE     Oxyluciferin, a luminescence product of firefly luciferase, is
            enzymatically regenerated into luciferin
  JOURNAL   J. Biol. Chem. 276 (39), 36508-36513 (2001)
  MEDLINE   21453302
REFERENCE   2  (bases 1 to 927)
  AUTHORS   Gomi,K. and Kajiyama,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2001) Keiko Gomi, Kikkoman Corporation, Research
            & Development; Noda 399, Noda-shi, Chiba 2780037, Japan
            (E-mail:8553@mail.kikkoman.co.jp, Tel:81471235579)
FEATURES             Location/Qualifiers
     source          1. .927
                     /organism="Photinus pyralis"
                     /db_xref="taxon:7054"
     CDS             1. .927
                     /codon_start=1
                     /product="luciferin regenerating enzyme"
                     /protein_id="BAB60700.1"
                     /db_xref="GI:14331152"
                     /translation="MGPVVEKIAELGKYTVGEGPHWDHETQTLYFVDTVEKTFHKYVP
                     SQKKYTFCKVDKLVSFIIPLAGSPGRFVVSLEREIAILTWDGVSAAPTSIEAIVNVEP
                     HIKNNRLNDGKADPLGNLWTGTMAIDAGLPVGPVTGSLYHLGADKKVKMHESNIAIAN
                     GLAWSNDLKKMYYIDSGKRRVDEYDYDASTLSISNQRPLFTFEKHEVPGYPDGQTIDE
                     EGNLWVAVFQGQRIIKISTQQPEVLLDTVKIPDPQVTSVAFGGPNLDELHVTSAGLQL
                     DDSSLDKSLVNGHVYRVTGLGVKGFAGVKVKL"
BASE COUNT      285 a    171 c    216 g    255 t
ORIGIN

  Query Match             99.9%;  Score 923;  DB 3;  Length 927;
  Best Local Similarity   99.9%;  Pred. No. 3e-220;
  Matches  923;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGGCCAGTTGTTGAAAAAATTGCAGAACTTGGCAAGTATACGGTTGGAGAAGGTCCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGGCCAGTTGTTGAAAAAATTGCAGAACTTGGCAAGTATACGGTTGGAGAAGGTCCT 60

Qy         61 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 120

Qy        121 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 180

Qy        181 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 240

Qy        241 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTC 300

Qy        301 GAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTA 360

Qy        361 TGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTA 420

Qy        421 TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAAT 480

Qy        481 GGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGA 540

Qy        541 GTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTT 600

Qy        601 ACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGT 660

Qy        661 AATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCG 720

Qy        721 GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGC 780

Qy        781 GGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGACAGTTCT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGACAGTTCT 840

Qy        841 TTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTCAAAGGT 900
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTGGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTCAAAGGT 900

Qy        901 TTCGCGGGAGTTAAAGTGAAGCTA 924
              |||||||||||||||||||||||||
Db        901 TTCGCGGGAGTTAAAGTGAAGCTA 924


RESULT 2
AB072447
LOCUS       AB072447                924 bp    mRNA    linear   INV 22-FEB-2002
DEFINITION  Luciola lateralis H-LRE mRNA for luciferin-regenerating enzyme,
            complete cds.
ACCESSION   AB072447
VERSION     AB072447.1  GI:18857904
KEYWORDS    .
SOURCE      Luciola lateralis cDNA to mRNA.
  ORGANISM  Luciola lateralis
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia;
            Cantharoidea; Lampyridae; Luciola.
REFERENCE   1
  AUTHORS   Gomi,K., Hirokawa,K. and Kajiyama,N.
  TITLE     Molecular cloning and expression of luciferin-regenerating enzyme
            (LRE) cDNA in Luciola cruciata and Luciola lateralis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 924)
  AUTHORS   Gomi,K., Hirokawa,K. and Kajiyama,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2001) Keiko Gomi, Kikkoman Corporation, Research
            and Development Division; Noda 399, Noda, Chiba 2780037, Japan
            (E-mail:8553@mail.kikkoman.co.jp, Tel:81471235579, Fax:81471235550)
FEATURES             Location/Qualifiers
     source          1. .924
                     /organism="Luciola lateralis"
                     /db_xref="taxon:7052"
     gene            1. .924
                     /gene="H-LRE"
     CDS             1. .924
                     /gene="H-LRE"
                     /codon_start=1
                     /product="luciferin-regenerating enzyme"
                     /protein_id="BAB85478.1"
                     /db_xref="GI:18857905"
                     /translation="MSPVIEQITEVDNFQIGEGPHWDTETQSLYFVDILEKSIHKYVP
                     STKQHTKMILNKRPSFIIPIKETSDRFVISLERDICVLTWDGVSATPSHLETIVTVDT
                     GIEGNTFNDGKADAFGNLWAGTLYSKFDIEKQGPNTGTLYSLSNKQLRKHISNIFLSN
                     GLAWNKDSKKFYFIDSNKRTIDQFDYDSENLIISNCQPLFTLDKHGIQGLPDAQTIDE
                     NDNLWVAIVRGGKVINIGTKQPESLLGVINMPESLITSVCFGGSKLDELYVTTSGIKE
                     YETDSTKLVKGGLYRVTGLGVKGLPAHRFSL"
BASE COUNT      325 a    146 c    178 g    275 t
ORIGIN

  Query Match             35.2%;  Score 325.2;  DB 3;  Length 924;
  Best Local Similarity   60.2%;  Pred. No. 1.1e-70;
  Matches  556;  Conservative    0;  Mismatches  364;  Indels    3;  Gaps    1;
```

```
Qy          1 ATGGGGCCAGTTGTTGAAAAAATTGCAGAACTTGGCAAGTATACGGTTGGAGAAGGTCCT 60
Db          1 ATGTCGCCAGTTATTGAACAGATCACTGAAGTAGACAATTTCCAAATCGGAGAGGGTCCA 60

Qy         61 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 120
Db         61 CACTGGGATACCGAAACACAAAGTTTGTATTTTGTGGATATTCTAGAAAAATCCATACAC 120

Qy        121 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 180
Db        121 AAATACGTACCATCGACAAAACAACATACTAAAATGATTTTGAATAAACGTCCGTCTTTT 180

Qy        181 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 240
Db        181 ATTATACCAATAAAAGAAACATCTGATCGGTTTGTCATAAGTTTAGAACGAGACATTTGC 240

Qy        241 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAAGCTATTGTTAATGTC 300
Db        241 GTTCTTACTTGGGATGGTGTTAGTGCCACGCCAAGTCATTTAGAAACAATAGTTACCGTT 300

Qy        301 GAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAATCTA 360
Db        301 GATACGGGAATCGAAGGAAATACATTCAATGATGGTAAAGCAGATGCGTTTGGCAATTTG 360

Qy        361 TGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTA 420
Db        361 TGGGCAGGTACATTATATAGTAAATTCGATATTGAAAAACAAGGTCCTAATACAGGAACA 420

Qy        421 TATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATAGCAAAT 480
Db        421 TTGTACAGCCTGTCTAATAAGCAGTTAAGAAAACATATTTCTAATATCTTCCTATCAAAT 480

Qy        481 GGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAAAGAAGA 540
Db        481 GGCCTAGCCTGGAATAAAGACTCAAAAAAATTTTATTTTATCGACTCCAATAAAAGAACA 540

Qy        541 GTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCATTATTT 600
Db        541 ATAGATCAGTTTGATTATGATTCTGAAAATTTAATTATATCAAATTGTCAACCATTGTTT 600

Qy        601 ACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGT 660
Db        601 ACTCTGGACAAACATGGGATACAGGGTTTACCTGATGCCCAAACGATAGATGAAAATGAT 660

Qy        661 AATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCG 720
Db        661 AATTTATGGGTTGCTATAGTTCGAGGAGGAAAAGTTATAAATATTGGTACTAAGCAACCC 720

Qy        721 GAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGC 780
Db        721 GAATCTTTGCTTGGTGTTATTAACATGCCTGAAAGTTTGATAACTTCAGTTTGTTTTGGA 780

Qy        781 GGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGACAGTTCT 840
Db        781 GGATCAAAGTTGGACGAACTTTATGTAACGACTTCTGGTATCAAAGAGTATGAAA---CT 837

Qy        841 TTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTCAAAGGT 900
Db        838 GACTCTACAAAACTAGTAAAGGGTGGATTGTACAGAGTTACTGGATTAGGTGTTAAAGGT 897

Qy        901 TTCGCGGGAGTTAAAGTGAAGCT 923
Db        898 TTACCCGCACACAGATTCAGTCT 920
```

```
RESULT 3
AB072448
LOCUS       AB072448                930 bp    mRNA    linear   INV 22-FEB-2002
DEFINITION  Luciola cruciata G-LRE mRNA for luciferin-regenerating enzyme,
            complete cds.
ACCESSION   AB072448
VERSION     AB072448.1  GI:18857906
KEYWORDS    .
SOURCE      Luciola cruciata cDNA to mRNA.
  ORGANISM  Luciola cruciata
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Coleoptera; Polyphaga; Elateriformia;
            Cantharoidea; Lampyridae; Luciola.
REFERENCE   1
  AUTHORS   Gomi,K., Hirokawa,K. and Kajiyama,N.
  TITLE     Molecular cloning and expression of luciferin-regenerating enzyme
            (LRE) cDNA in Luciola cruciata and Luciola lateralis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Gomi,K., Hirokawa,K. and Kajiyama,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2001) Keiko Gomi, Kikkoman Corporation, Research
            and Development Division; Noda 399, Noda, Chiba 2780037, Japan
            (E-mail:8553@mail.kikkoman.co.jp, Tel:81471235579, Fax:81471235550)
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Luciola cruciata"
                     /db_xref="taxon:7051"
     gene            1. .930
                     /gene="G-LRE"
     CDS             1. .930
                     /gene="G-LRE"
                     /codon_start=1
                     /product="luciferin-regenerating enzyme"
                     /protein_id="BAB85479.1"
                     /db_xref="GI:18857907"
                     /translation="MAPTVEQIVELGTYLLAESPHWDDETQSLYFVDIVGRSVNKYVP
                     TTKTHTQLKFDKNPSFIIPVKGCSDRFIVSLEREINLLTWDGASSAPSKIEKIAVFDN
                     TPEKSENRLNDGKADPLGNLWAGTMNMGSDHTTGTPVRGTLSSLSNKQVKEHVSEVCI
                     SNGLAWSKDLKKFYYIDSAVRQVDQFDFDAKNLSLSNRQPLFTFDKHGIMGSPDGQTI
                     DAEGNLWVATCQGDKVLKIDTSTPETLLGIVEIPEHQVTSVCIGGAELNVLYVTTASI
                     KLPGADETKPMKGAIYKVTGLGVKGLPGDRVKL"
BASE COUNT      305 a    150 c    203 g    272 t
ORIGIN
```

```
  Query Match             34.7%;  Score 320.4;  DB 3;  Length 930;
  Best Local Similarity   60.6%;  Pred. No. 1.7e-69;
  Matches  563;  Conservative    0;  Mismatches  357;  Indels    9;  Gaps    2;

Qy          1 ATGGGGCCAGTTGTTGAAAAAATTGCAGAACTTGGCAAGTATACGGTTGGAGAAGGTCCT 60
Db          1 ATGGCTCCAACTGTTGAACAAATAGTGGAATTGGGCACTTATTTGCTTGCAGAAAGTCCA 60

Qy         61 CACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAGAGAAAACTTTTCAT 120
Db         61 CATTGGGACGACGAAACTCAAAGTTTGTACTTTGTGGATATTGTAGGAAGATCTGTAAAC 120

Qy        121 AAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAAGTAGATAAACTGGTTTCTTTC 180
Db        121 AAGTACGTGCCGACTACCAAAACGCACACACAACTGAAATTTGATAAAAATCCATCGTTT 180

Qy        181 ATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCAGTTTGGAACGTGAAATAGCC 240
Db        181 ATCATTCCTGTCAAAGGATGTTCCGATCGTTTTATTGTGAGTTTAGAACGAGAAATTAAT 240

Qy        241 ATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCATAGAA------GCTATTGTT 294
Db        241 CTTCTTACATGGGATGGTGCTAGTTCTGCTCCAAGTAAGATAGAAAAAATCGCTGTATTT 300

Qy        295 AATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGC 354
Db        301 GATAATACTCCTGAAAAAAGTGAAAATAGATTAAATGACGGTAAAGCAGACCCTCTTGGA 360

Qy        355 AATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAGGACCGGTCACTGGC 414
Db        361 AATCTGTGGGCTGGAACGATGAATATGGGTTCAGATCATACGACAGGAACACCCGTTCGT 420

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
Db        421 GGCACGTTGTCGAGTTTATCTAATAAGCAAGTGAAAGAACACGTGTCTGAAGTTTGTATA 480
```

```
Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
Db        481 TCCAATGGTCTTGCTTGGAGTAAAGATTTAAAAAAGTTTTATTACATTGATTCTGCTGTT 540

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
Db        541 AGACAAGTCGATCAATTTGATTTTGATGCAAAAAATTTATCACTTTCTAACCGACAACCG 600

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
Db        601 TTATTTACATTTGACAAACATGGAATTATGGGTTCACCTGATGGACAGACTATAGATGCC 660

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
Db        661 GAAGGAAACTTATGGGTAGCCACATGTCAAGGTGATAAAGTTTTAAAAATTGATACTAGT 720

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
Db        721 ACTCCGGAAACCTTACTTGGAATTGTCGAGATTCCAGAACATCAGGTGACATCAGTTTGC 780

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
Db        781 ATTGGTGGAGCGGAATTAAATGTATTGTATGTTACTACTGCCAGTATAAAACTTCCCG-- 838

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
Db        839 -GTGCTGATGAAACAAAACCAATGAAGGGTGCTATCTATAAAGTAACTGGACTTGGTGTG 897

Qy        895 AAAGGTTTCGCGGGAGTTAAAGTGAAGCT 923
Db        898 AAAGGATTGCCAGGAGATCGAGTTAAGTT 926
```

```
RESULT 4
AY071130
LOCUS       AY071130                1040 bp    mRNA    linear   INV 20-DEC-2001
DEFINITION  Drosophila melanogaster RE17477 full length cDNA.
ACCESSION   AY071130
VERSION     AY071130.1  GI:17945393
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1040)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,
            George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,
            Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Patel,S., Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
            and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
            This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone. However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .1040
                     /organism="Drosophila melanogaster"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /map="88D2-88D2"
     gene            1. .1040
                     /gene="smp-30"
                     /note="alignment with genomic scaffold AE003706"
                     /db_xref="FLYBASE:FBgn0038257"
     CDS             49. .960
                     /gene="smp-30"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="RE17477p"
                     /protein_id="AAL48752.1"
                     /db_xref="GI:17945394"
                     /db_xref="FLYBASE:FBgn0038257"
                     /translation="MSYKVEAVPDSYAALGEGPHWDVDRQSLYYVDLESAGINRYDFK
                     QNKVYRAKIEGEIFASFILPVENKPQEFAVGCGLRTVIVQWDGVSAVAKVTRTLFEVQ
                     PDLKENRLNDAKTDPNGRFYGGTMADSGDIFTQWKGELYSWQAGGQPNAIRSKVGISN
                     GLAWDVKAKKFYFIDTNNHEVLAYDYNQSTGAVSNPKVIFDLRKIRPEGPLFPDGMTV
                     DTDGNIYVATFNGGTVFKVNPSTGKILLEIKIPTTQITSVAFGGPNLDILYVTTANKF
                     DQPKPAGTTFQVTGLNAKGYAGVNLKI"
BASE COUNT      287 a    262 c    254 g    237 t
ORIGIN
```

```
  Query Match             10.7%;  Score 98.6;  DB 3;  Length 1040;
  Best Local Similarity   49.2%;  Pred. No. 5.2e-14;
  Matches  384;  Conservative    0;  Mismatches  379;  Indels   18;  Gaps    4;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
Db         92 TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 151

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAA------G 160
Db        152 CCGCCGGCATTAATCGTTATGATTTCAAGCAGAACAAAGTGTACAGGGCTAAAATCGAGG 211

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
Db        212 GCGAGATATTTGCATCGTTCATTCTGCCGGTTGAGAACAAACCGCAGGAGTTTGCCGTAG 271

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
Db        272 GATGCGGTCTTCGTACGGTCATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAGGTCA 331

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
Db        332 CTCGCACCCTGTTCGAGGTGCAGCCGGACCTGAAGGAAAACCGCCTTAATGATGCCAAAA 391

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
Db        392 CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGACATATTCA 448

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
Db        449 CCCAATGGAAGGGTGAGCTCTACAGCTGGCAGGCCGGTGGACAGCCCAACGCTATCCGTA 508

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
Db        509 GCAAGGTGGGCATATCCAATGGCCTGGCCTGGGATGTCAAGGCCAAGAAGTTCTACTTCA 568

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
Db        569 TCGACACCAACAACCACGAGGTATTGGCCTATGACTACAATCAGAGCACCGGCGCCGTAA 628

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAA------GCATGAAGTGCCTGGATATCCAG 634
Db        629 GCAACCCAAAGGTCATCTTCGATCTGAGGAAGATTCGGCCCGAAGGACCATTGTTCCCTG 688
```

```
Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
              ||||    ||  | ||    || || ||| | |  || |||   ||| | ||
Db        689 ATGGCATGACCGTAGACACCGATGGCAATATCTACGTGGCCACCTTCAATGGTGGCACCG 748

Qy        695 TTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATC 754
              |  | ||  ||| | |           || |  | ||||| |   | ||||| |||
Db        749 TCTTCAAGGTCAATCCAAGCACCGGTAAAATTCTGCTGGAGA---TCAAAATTCCAACCA 805

Qy        755 CTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTG 814
              | ||  ||||||| || || ||||| ||||| |||||||||    || |||| ||| | |
Db        806 CCCAAATCACCTCGGTGGCTTTTGGAGGTCCCAATTTGGATATTTTGTATGTGACAACCG 865

Qy        815 C 815
              |
Db        866 C 866
```

```
RESULT 5
AY028616
LOCUS       AY028616               1150 bp    mRNA    linear   INV 29-MAR-2001
DEFINITION  Calliphora vicina anterior fat body protein mRNA, complete cds.
ACCESSION   AY028616
VERSION     AY028616.1  GI:13488604
KEYWORDS    .
SOURCE      Calliphora vicina.
  ORGANISM  Calliphora vicina
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Oestroidea; Calliphoridae; Calliphora.
REFERENCE   1  (bases 1 to 1150)
  AUTHORS   Hansen,I.A., Meyer,S.R., Schaefer,I. and Scheller,K.
  TITLE     Identification of the anterior fat body protein as interactor of
            the hexamerin receptor in the blowfly Calliphora vicina
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1150)
  AUTHORS   Hansen,I.A., Meyer,S.R., Schaefer,I. and Scheller,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-2001) Cell and Developmental Biology, University
            Wuerzburg, Biocenter Hubland, Wuerzburg D-97974, Germany
FEATURES             Location/Qualifiers
     source          1. .1150
                     /organism="Calliphora vicina"
                     /db_xref="taxon:7373"
                     /cell_type="trophocytes"
                     /tissue_type="anterior fat body"
                     /dev_stage="larvae"
     CDS             42. .962
                     /note="AFP; strong interactor with the hexamerin receptor
                     in a yeast-2-hybrid interaction screen; CvAFP"
                     /codon_start=1
                     /product="anterior fat body protein"
                     /protein_id="AAK26174.1"
                     /db_xref="GI:13488605"
                     /translation="MSYKVEPLPDSHAELGEGPHWDVATQNLYYVDINAGKLLRYNYN
                     ENKVYKAKIEGEDLAGFAVPVEGTTDQFAVGSGRRVVIVQWDGVSETAKVLKTLFEVQ
                     TGDERFTGNRFNDGKCDPRGRLFAGTMRYVGDEFEHRYGELYKYENGGEVEVIKSDVG
                     ISNGLAWNEKTKKFYYIDTTDYEVKEYDYDFETGKASNPKVVFNLRKTSPKDHLLPDG
                     MTIDTEGNLYVATFNGHSIYKINPTTGQVLLEIKFPCKQITSAAFGGPNFHILFVTTS
                     SRFGEPHPAGTTYKVTGLGAKGYPMTKIQL"
BASE COUNT      338 a    223 c    231 g    358 t
ORIGIN
```

```
  Query Match           10.5%;  Score 97.2;  DB 3;  Length 1150;
  Best Local Similarity   50.9%;  Pred. No. 1.2e-13;
  Matches  399;  Conservative    0;  Mismatches  358;  Indels   27;  Gaps    6;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              | || ||||| ||||| ||||||  ||  || ||||  || || |  || |||| |   |
Db         85 TAGGCGAAGGACCTCATTGGGATGTTGCCACCCAGAATTTGTACTATGTTGACATCAATG 144

Qy        107 --AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAA----AATACACGTTTTGTAAAG 160
                | ||||| ||     | |  ||| |   |  |||    | |  |      | |||
Db        145 CCGGCAAACTGTTGCGTTACAACTACAATGAAAATAAAGTGTACAAGGCCAAAATTGAAG 204

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
                ||  |  | | |  |||   | | ||| | |  ||  |  | |  |  |||| |||
Db        205 GTGAGGATTTAGCTGGTTTTGCTGTGCCCGTCGAAGGTACAACCGATCAATTTGCTGTCG 264

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              |||  |  |||    | | |||| |   |||||||||| |||  ||   |  ||| |
Db        265 GTTCTGGTCGTCGTGTGGTCATTGTACAATGGGATGGTGTTTCTGAGACCGCTAAAGTAT 324

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAAT---------AACAGACTCAATG 331
              |  |  || |  |  | ||  || |      | ||  |           ||| |  ||||||
Db        325 TGAAGACTTTGTTCGAAGTGCAAACCGGTGATGAACGTTTCACTGGCAACCGTTTCAATG 384

Qy        332 ATGGCAAAGCAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTC 391
              |||||||    ||||| | |||   | | |   | ||||| |||  |  || ||  | |
Db        385 ATGGCAAGTGTGATCCTCGTGGTCGTTTGTTTGCCGGTACCATGCGTTATGTCGGCGATG 444

Qy        392 TCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAA 451
                ||     |      |||   |||||  | |  |   ||      | ||  |
Db        445 AATTCGAGCATCGTTA---TGGTGAATTATACAAGTACGAAAATGGAGGCGAAGTTGAGG 501

Qy        452 TGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAA 511
              |   | |     |  | | |||  | |||||  | || |||| | |  |     ||||||
Db        502 TAATCAAATCGGATGTGGGTATTTCCAATGGTTTGGCTTGGAATGAGAAGACTAAGAAAT 561

Qy        512 TGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACAT 571
              | ||||| || ||| |     |       ||  | ||||| |||||||||  |   ||
Db        562 TCTATTACATCGATACCACCGATTATGAGGTTAAGGAGTATGATTATGATTTTGAAACTG 621

Qy        572 TATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAG------CATGAAGTGCCTG 625
                  |  |||||         |||| | | |    |||      |   | | |  | |
Db        622 GCAAGGCCTCCAATCCCAAAGTTGTATTCAATCTACGCAAGACCAGCCCCAAGGATCATT 681

Qy        626 GATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAG 685
              |  || |||||    || ||||||  |||||  ||| | |  || |||  |||| | |
Db        682 TGTTGCCCGATGGCATGACCATTGATACTGAGGGCAATCTTTATGTGGCCACTTTCAATG 741

Qy        686 GACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAA 745
              | ||    ||    ||||| | | ||          |||| ||  |||| |   | |||
Db        742 GCCATTCCATCTACAAAATTAATCCCACCACTGGCCAAGTTTTGTTGGAAA---TCAAAT 798

Qy        746 TACCAGATCCTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATG 805
              | ||        ||  |||||||  |  || ||||| ||||| |||||   ||  ||  ||
Db        799 TCCCCTGCAAACAAATCACCTCGGCTGCTTTTGGTGGTCCCAATTTCCATATTTTGTTTG 858

Qy        806 TAAC 809
              ||||
Db        859 TAAC 862
```

```
RESULT 6
AB029490
LOCUS       AB029490               1022 bp    mRNA    linear   INV 03-MAY-2002
DEFINITION  Drosophila melanogaster mRNA for senescence marker protein-30
            (SMP-30), complete cds.
ACCESSION   AB029490
VERSION     AB029490.1  GI:7339661
KEYWORDS    senescence marker protein-30 (SMP-30).
SOURCE      Drosophila melanogaster (strain:Canton S) 7 day adult stage whole
            body cDNA to mRNA.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1
  AUTHORS   Goto,S.G.
  TITLE     Expression of Drosophila homologue of senescence marker protein-30
            during cold acclimation
  JOURNAL   J. Insect Physiol. 46 (7), 1111-1120 (2000)
   PUBMED   10817837
REFERENCE   2  (bases 1 to 1022)
  AUTHORS   Goto,S.G.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (01-JUL-1999) Shin G Goto, Ohio State University,
            Department of Entomology; 1735 Neil Avenue, Columbus, OH 43210, USA
            (E-mail:shingoto@sci.osaka-cu.ac.jp, Tel:+1-614-292-4477,
            Fax:+1-614-292-2180)
FEATURES             Location/Qualifiers
     source          1. .1022
                     /organism="Drosophila melanogaster"
                     /strain="Canton S"
                     /db_xref="taxon:7227"
                     /tissue_type="whole body"
                     /dev_stage="7 day adult stage"
     CDS             29. .940
                     /codon_start=1
                     /product="senescence marker protein-30 (SMP-30)"
                     /protein_id="BAA92938.1"
                     /db_xref="GI:7339662"
                     /translation="MSYKVEAVPDSYAALGEGPHWDVDRQSLYYVDLESAGINRYDFK
                     QNKVYKAKIEGEIFASFILPVENKPQEFAVGCGLRTVIVQWDGVSAVAKVTRTLFEVQ
                     PDLKENRLNDAKTDPNGRFYGGTMADSGDIFTQWKGELYSWQAGGQPNAIRSKVGISN
                     GLAWDVKAKKFYFIDTNNHEVLAYDYNQSTGAVSNPKVIFDLRKIRPEGPLFPDGMTV
                     DTDGNIYVATFNGGTVFKVNPSTGKILLEIKIPTTQITSVAFGGPNLDILYVTTANKF
                     DQPKPAGTTFQVTGLNAKGYAGVNLKI"
BASE COUNT      287 a    260 c    249 g    226 t
ORIGIN

  Query Match             10.2%;  Score 93.8;  DB 3;  Length 1022;
  Best Local Similarity   48.8%;  Pred. No. 8.3e-13;
  Matches  381;  Conservative    0;  Mismatches  382;  Indels   18;  Gaps    4;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              | || || || || |||||||||  |||    ||||   | || | ||| |||  || |
Db         72 TGGGCGAGGGACCCCACTGGGATGTTGATCGCCAGAGTCTGTACTACGTGGACCTCGAAT 131

Qy        107 AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAAAATACACGTTTTGTAAA------G 160
                  || ||   || |        ||||| |||           ||||      |
Db        132 CCGCCGGCATTAATCGTTACGATTTCAAGCAGAACAAAGTGTACAAGGCTAAAATCGAGG 191

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
              || | | | |  || ||||| | ||  |||        ||       ||||  ||
Db        192 GCGAGATATTTGCATCCTTCATTCTGCCGGTTGAGAACAAACCGCAGGAGTTTGCCGTAG 251

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              | |  |  | |   | |||  |    ||||||| ||    || |   | |    ||
Db        252 GATGCGGTCTTCGTACGGTCATCGTCCAGTGGGATGGAGTCTCCGCAGTGGCCAAGGTCA 311

Qy        281 TAGAAGCTATTGTTAATGTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAG 340
                | |  | | ||  | ||  |  | ||  | ||| | || |||||||  ||||
Db        312 CTCGCACCCTGTTCGAGGTGCAGCCGGATCTGAAGGAAAACCGCCTTAATGATGCCAAAA 371

Qy        341 CAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTAG 400
              | ||||||  ||||  | | |     || || |||||    |||  |||      | |
Db        372 CCGATCCCAATGGCCGTTTTTACGGTGGCACCATGGC---CGACAGTGGCGATATCTTCA 428

Qy        401 GACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGA 460
              |     |  ||     | ||    |    ||| | |  | ||   ||       |   |
Db        429 CCCAATGGAAGGGTGAGCTCTACAGCTGGCAGGCCGGTGGACAGCCCAACGCTATCCGTA 488

Qy        461 GCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATA 520
              ||||  | |  ||| | ||||| || || |||  |    |    |||||  | || |  |
Db        489 GCAAGGTGGGCATATCCAATGGCCTGGCCTGGGATGTCAAGGCCAAGAAGTTCTACTTCA 548

Qy        521 TTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCA 580
              | ||  |    ||        |||   |  || || ||  ||        ||   || |
Db        549 TCGACACCAACAACCACGAGGTATTGGCCTATGACTACAATCAGAGCACCGGCGCCGTGA 608

Qy        581 GCAATCAACGGCCATTATTTACTTTTGAAAA------GCATGAAGTGCCTGGATATCCAG 634
              |||| | |   |  | ||   | |    ||      ||  ||||  ||    |  || |
Db        609 GCAACCCAAAGGTCATCTTCGATCTGAGGAAGATTCGGCCCGAAGGACCATTGTTCCCTG 668

Qy        635 ATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAA 694
              ||||    ||  | ||    || || ||| | |  || |||   ||| | ||
Db        669 ATGGCATGACCGTAGACACCGATGGCAATATCTACGTGGCCACCTTCAATGGTGGCACCG 728

Qy        695 TTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATC 754
              |  | ||  ||| | |           || |  | ||||| |   | ||||| |||
Db        729 TCTTCAAGGTCAATCCAAGCACCGGTAAAATCCTGCTGGAGA---TCAAAATTCCAACCA 785

Qy        755 CTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTG 814
              | ||  ||||||| || || ||||| ||||| || ||||||    || |||| ||| | |
Db        786 CCCAAATCACCTCGGTGGCTTTTGGAGGTCCCAACTTGGATATTTTGTATGTGACAACCG 845

Qy        815 C 815
              |
Db        846 C 846


RESULT 7
AF326959
LOCUS       AF326959               992 bp    mRNA    linear   INV 08-MAY-2001
DEFINITION  Calliphora vicina anterior fat body protein (AFP) mRNA, partial
            cds.
ACCESSION   AF326959
VERSION     AF326959.1  GI:13991404
KEYWORDS    .
SOURCE      Calliphora vicina.
  ORGANISM  Calliphora vicina
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Oestroidea; Calliphoridae; Calliphora.
REFERENCE   1  (bases 1 to 992)
  AUTHORS   Hansen,I.A., Meyer,S.R., Berlinger,M.J. and Scheller,K.
  TITLE     Identification of new interactors of the hexamerin receptor in the
            blowfly Calliphora vicina
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 992)
  AUTHORS   Hansen,I.A., Meyer,S.R., Berlinger,M.J. and Scheller,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-DEC-2000) Zoologie 1, University of Wuerzburg,
            Biocenter Hubland, Wuerzburg 97074, Germany
FEATURES             Location/Qualifiers
     source          1. .992
                     /organism="Calliphora vicina"
                     /db_xref="taxon:7373"
                     /tissue_type="fat body"
                     /dev_stage="larvae"
     gene            <1. .992
                     /gene="AFP"
     CDS             <1. .804
                     /gene="AFP"
                     /note="found as strong interactor of the hexamerin
                     receptor in a yeast 2-hybrid screen"
                     /codon_start=1
                     /product="anterior fat body protein"
                     /protein_id="AAK51353.1"
                     /db_xref="GI:13991405"
                     /translation="GYNYNENKVYKAKIEGEDLAGFAVPVEGTTDQFAVGSGRRVVIV
                     QWDGVSETAKVLKTLFEVQTGDERFTGNRFNDGKCDPRGRLFAGTMRYVGDEFEHRYG
                     ELYKYENGGEVEVIKSDVGISNGLAWNEKTKKFYYIDTTDYEVKEYDYDFETGKASNP
                     KVVFNLRKTSPKDHLLPDGMTIDTEGNLYVATFNGHSIYKINPTTGQVLLEIKFPCKQ
                     ITSAAFGGPNFHILFVTTSSRFGEPHPAGTTYKVTGLGAKGYPMTKIQL"
BASE COUNT      293 a    189 c    201 g    309 t
ORIGIN

  Query Match             9.8%;  Score 91;  DB 3;  Length 992;
  Best Local Similarity   50.5%;  Pred. No. 4.2e-12;
  Matches  338;  Conservative    0;  Mismatches  310;  Indels   21;  Gaps    4;

Qy        156 TAAAGTAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGT 215
              | |||  ||  |  | |'|  |||   | | ||| | |  ||  |  | |  ||||
Db         42 TGAAGGTGAGGATTTAGCTGGTTTTGCTGTGCCCGTCGAAGGTACAACCGATCAATTTGC 101

Qy        216 AGTCAGTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTAC 275
                |||  |||  |  |||    | | |||| |   |||||||||  |||  ||   |  |||
Db        102 TGTCGGTTCTGGTCGTCGTGTGGTCATTGTACAATGGGATGGTGTTTCTGAGACCGCTAA 161

Qy        276 AAGCATAGAAGCTATTGTTAATGTCGAACCACACATTAAAAAT---------AACAGACT 326
```

```
Db        162 AGTATTGAAGACTTTGTTCGAAGTGCAAACCGGTGATGAACGTTTCACTGGCAACCGTTT 221

Qy        327 CAATGATGGCAAAGCAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGC 386

Db        222 CAATGATGGCAAGTGTGATCCTCGTGGTCGTTTGTTTGCCGGTACCATGCGTTATGTCGG 281

Qy        387 TGGTCTCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGT 446

Db        282 CGATGAATTCGAGCATCGTTA---TGGTGAATTATACAAGTACGAAAATGGAGGCGAAGT 338

Qy        447 AAAAATGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAA 506

Db        339 TGAGGTAATCAAATCGGATGTGGGTATTTCCAATGGTTTGGCTTGGAATGAGAAGACTAA 398

Qy        507 GAAAATGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTC 566

Db        399 GAAATTCTATTACATCGATACCACCGATTATGAGGTTAAGGAGTATGATTATGATTTTGA 458

Qy        567 TACATTATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAG------CATGAAGT 620

Db        459 AACTGGCAAGGCCTCCAATCCCAAAGTTGTATTCAATCTACGCAAGACCAGCCCCAAGGA 518

Qy        621 GCCTGGATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTT 680

Db        519 TCATTTGTTGCCCGATGGCATGACCATTGATACTGAGGGCAATCTTTATGTGGCCACTTT 578

Qy        681 CCAAGGACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGT 740

Db        579 CAATGGCCATTCCATCTACAAAATTAATCCCACCACTGGCCAAGTTTTGTTGGAAATC-- 636

Qy        741 AAAAATACCAGATCCTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACT 800

Db        637 -AAATTCCCCTGCAAACAAATCACCTCGGCTGCTTTTGGTGGTCCCAATTTCCATATTTT 695

Qy        801 GCATGTAAC 809

Db        696 GTTTGTAAC 704
```

```
RESULT 8
E09237
LOCUS       E09237                  1216 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  cDNA insert encoding Regucalcin.
ACCESSION   E09237
VERSION     E09237.1  GI:22025863
KEYWORDS    JP 1995123985-A/1.
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1216)
  AUTHORS   Yamaguchi,M.
  TITLE     DNA FRAGMENT CODING FOR REGUCALCIN
  JOURNAL   Patent: JP 1995123985-A 1 16-MAY-1995;
            YAMAGUCHI MASAYOSHI, DAI ICHI PURE CHEM CO LTD
COMMENT     OS   Rattus sp. (Rat)
            PN   JP 1995123985-A/1
            PD   16-MAY-1995
            PF   09-NOV-1993 JP 1993279349
            PI   YAMAGUCHI MASAYOSHI
            PC   C12N15/09,C12N1/21,(C12N1/21,C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .1216
            FT                   /organism='Rattus sp.'
            FT   CDS             80. .979
            FT                   /product='Regucalcin'.
FEATURES             Location/Qualifiers
     source          1. .1216
                     /organism="Rattus sp."
                     /db_xref="taxon:10118"
BASE COUNT      338 a    231 c    309 g    338 t
ORIGIN
```

```
  Query Match           9.8%;  Score 90.4;  DB 6;  Length 1216;
  Best Local Similarity   50.2%;  Pred. No. 5.8e-12;
  Matches  307;  Conservative    0;  Mismatches  292;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357

Db        356 GTGGATCAAGATAAGAAAAACAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 415

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414

Db        416 TACTTTGCTGGTACCATGGCTGAGGAAACCGCCCCAGCTGTTCTGGAGCGGCACCAAGGG 475

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474

Db        476 TCCTTGTACTCCCTTTTTCCTGATGACAGTGTGAAGAAATACTTTAACCAAGTGGATATC 535

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534

Db        536 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 595

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594

Db        596 TACACTGTGGATGCCTTTGACTATGACCTGCCAACAGGACAGATTTCCAACCGCAGGACT 655

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654

Db        656 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGTT 709

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714

Db        710 GAGGGGAAGCTTTGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTAGATCCTGAG 769

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774

Db        770 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 826

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834

Db        827 TTTGGAGGGAAGGATTACTCTGAAATGTACGTGACATGTGCCAGGGATGGGATGAGCGCC 886

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894

Db        887 GAAGGTCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTTGGGGTC 946

Qy        895 AAAGGTTTCGC 905

Db        947 AAAGGAATTGC 957
```

```
RESULT 9
AB036903
LOCUS       AB036903                1155 bp    mRNA    linear   INV 12-JUL-2000
DEFINITION  Sarcophaga peregrina AFP mRNA for anterior fat body protein,
            complete cds.
ACCESSION   AB036903
VERSION     AB036903.1  GI:8980620
KEYWORDS    AFP; anterior fat body protein.
SOURCE      Sarcophaga peregrina larva fat body cDNA to mRNA.
  ORGANISM  Sarcophaga peregrina
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Oestroidea; Sarcophagidae; Sarcophaga.
REFERENCE   1  (sites)
  AUTHORS   Nakajima,Y. and Natori,S.
  TITLE     Identification and characterization of an anterior fat body protein
            in an insect
  JOURNAL   J. Biochem. 127 (5), 901-908 (2000)
  MEDLINE   20250851
```

```
REFERENCE   2  (bases 1 to 1155)
  AUTHORS   Nakajima,Y. and Natori,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JAN-2000) Yuki Nakajima, Riken, Natori Special Lab.;
            Hirosawa2-1, Wako, Saitama 351-0198, Japan
            (E-mail:nyuki@postman.riken.go.jp, Tel:+81-048-467-9439,
            Fax:+81-048-462-4693)
FEATURES             Location/Qualifiers
     source          1. .1155
                     /organism="Sarcophaga peregrina"
                     /db_xref="taxon:7386"
                     /tissue_type="fat body"
                     /dev_stage="larva"
     gene            1. .1155
                     /gene="AFP"
     CDS             86. .1006
                     /gene="AFP"
                     /codon_start=1
                     /product="anterior fat body protein"
                     /protein_id="BAA99282.1"
                     /db_xref="GI:8980621"
                     /translation="MSYKVEPLPDSHAELGEGPHWDVATQNLYYVDINAGKLLRYNYN
                     ENKVYKAKIEGEDLAGFAIPVEGTTDQFAVGSGRRVVIVQWDGISETAKVIKTLFEVQ
                     TGDERFNGNRFNDGKCDPRGRLFAGTMRYVGDEFEHRYGELYKYEKGGKIEVVKSNVG
                     ISNGLAWNEKTKKFYYIDTTDYEVKEYDYDFETGKATNPKVVFNLRKNSPKDHLLPDG
                     MTIDTDGNIYVATFNGYTIYKVNPTTGKVLLEIKFPCKQITSAAFGGPNLDILYVTTS
                     SRFGEPDPAGTTYKVTGLGAKGYPMTKIQL"
BASE COUNT      350 a    210 c    222 g    373 t
ORIGIN

  Query Match             9.7%;  Score 89.8;  DB 3;  Length 1155;
  Best Local Similarity   49.8%;  Pred. No. 8.2e-12;
  Matches  406;  Conservative    0;  Mismatches  382;  Indels   27;  Gaps    6;

Qy         47 TTGGAGAAGGTCCTCACTGGGATCATGAAACTCAGACCTTATATTTCGTCGACACCGTAG 106
              |||| ||||| || || |||||| ||  || || |  || |||| || ||||     |
Db        129 TTGGTGAAGGCCCACATTGGGATGTTGCTACACAAAATTTGTATTATGTTGACATTAATG 188

Qy        107 --AGAAAACTTTTCATAAATATGTACCTTCTCAGAAAA----AATACACGTTTTGTAAAG 160
                | ||||| ||      | |  ||| |   |  |||     |  |  |       | |||
Db        189 CTGGCAAACTCTTACGTTACAACTACAATGAAAATAAAGTGTACAAAGCAAAAATTGAAG 248

Qy        161 TAGATAAACTGGTTTCTTTCATTATTCCCCTTGCTGGATCCCCTGGCCGTTTTGTAGTCA 220
               ||  |  | | |  |||   ||| ||  | |  ||  |  | |  |  ||||  |||
Db        249 GTGAGGATTTAGCTGGTTTTGCTATACCTGTCGAAGGTACAACCGATCAGTTTGCCGTCG 308

Qy        221 GTTTGGAACGTGAAATAGCCATTCTTACATGGGATGGCGTTAGTGCTGCACCTACAAGCA 280
              |||  |  |||    | |  ||  |  |||||||||  |   ||   |  | | |  ||
Db        309 GTTCTGGTCGTCGTGTTGTTATCGTCCAATGGGATGGTATCTCTGAAACGGCCAAAGTCA 368

Qy        281 TAGAAGCTATTGTTAATGTCGAACC---------ACACATTAAAAATAACAGACTCAATG 331
              |  |  || |  || | ||  || |          ||. ||||   |||  |  | || |
Db        369 TTAAGACTTTGTTTGAAGTGCAAACCGGTGATGAACGTTTTAACGGTAATCGTTTTAACG 428

Qy        332 ATGGCAAAGCAGATCCCCTTGGCAATCTATGGACAGGTACAATGGCTATTGACGCTGGTC 391
              |||| |||   ||||| | |||   | | |   | |||| ||| |    |||  |||
Db        429 ATGGTAAATGTGATCCTCGTGGACGTTTGTTCGCTGGTACTATGCGTT---ACGTAGGTG 485

Qy        392 TCCCCGTAGGACCGGTCACTGGCAGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAA 451
              |    | | ||       |||    || ||  | |  |     | |   ||  |  |
Db        486 ACGAATTCGAACATCGTTATGGTGAATTGTACAAATACGAAAAAGGTGGCAAAATTGAGG 545

Qy        452 TGCACGAGAGCAACATAGCTATAGCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAA 511
              |     |    ||| | | |||  | ||||| ||||| |||| | |  |    ||||||
Db        546 TAGTTAAATCTAACGTGGGTATTTCTAATGGTCTCGCTTGGAATGAAAAGACAAAGAAAT 605

Qy        512 TGTATTATATTGATTCGGGGAAAAGAAGAGTAGACGAGTACGATTATGATGCTTCTACAT 571
              | ||||| |||||| |.    |        ||  | || || ||||| ||   |   ||
Db        606 TCTATTACATTGATACCACTGACTATGAGGTCAAAGAATATGATTACGACTTTGAGACTG 665

Qy        572 TATCCATCAGCAATCAACGGCCATTATTTACTTTTGAAAAGCAT------GAAGTGCCTG 625
                  |  ||||    |    | || | | |    ||  ||       | |  | |
Db        666 GCAAGGCTACAAATCCCAAGGTTGTTTTCAATCTACGCAAAAATAGTCCCAAGGATCATC 725

Qy        626 GATATCCAGATGGTCAAACAATTGATGAGGAGGGTAATTTATGGGTTGCCGTTTTCCAAG 685
              ||  || || |||   || |||||    || || ||| | |  || |||   ||  | |
Db        726 TATTGCCCGACGGTATGACCATTGACACTGATGGCAATATTTATGTAGCCACATTTAACG 785

Qy        686 GACAGCGAATTATTAAAATCAGTACCCAACAACCGGAAGTGTTACTGGATACCGTAAAAA 745
              |  |    ||    ||| | | | ||          |||| ||  |||| |   | |||
Db        786 GTTATACCATCTACAAAGTTAATCCCACTACCGGTAAAGTTTTGTTGGAAA---TTAAAT 842

Qy        746 TACCAGATCCTCAGGTCACCTCTGTAGCATTTGGCGGTCCGAATTTGGATGAACTGCATG 805
              | ||        ||| | ||||| |  ||||||||||||||| ||| | ||     || |||
Db        843 TCCCTTGCAAACAGATAACCTCAGCTGCATTTGGCGGTCCAAATCTAGACATTTTGTATG 902

Qy        806 TAACATCTGCTGGTCTTCAGCTTGACGACAGTTCT 840
              | ||| |  ||   | |      |||    |  | ||
Db        903 TCACAACATCTTCCCGTTTTGGTGAACCCGATCCT 937


RESULT 10
RATCBP1
LOCUS       RATCBP1                 1594 bp    mRNA    linear   ROD 08-FEB-1999
DEFINITION  Rat mRNA for calcium-binding protein, complete cds.
ACCESSION   D38467 D14327 D16386
VERSION     D38467.1  GI:600378
KEYWORDS    calcium-binding protein; regucalcin.
SOURCE      Rattus norvegicus (strain:Wistar) liver cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Shimokawa,N. and Yamaguchi,M.
  TITLE     Molecular cloning and sequencing of the cDNA coding for a
            calcium-binding protein regucalcin from rat liver
  JOURNAL   FEBS Lett. 327 (3), 251-255 (1993)
  MEDLINE   93351639
REFERENCE   2  (bases 1 to 1594)
  AUTHORS   Shimokawa,N. and Yamaguchi,M.
  JOURNAL   Unpublished
COMMENT     On Dec 8, 1994 this sequence version replaced gi:391851.
            D16386:Submitted(07-Jun-1993)to DDBJ by:Masayoshi Yamaguchi
            Submitted (03-OCT-1994) to DDBJ by:
            Masayoshi Yamaguchi
            Graduate School of Nutritional Sciences
            University of Shizuoka
            52-1 Yada
            Shizuoka 422.
FEATURES             Location/Qualifiers
     source          1. .1594
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /tissue_type="liver"
     CDS             93. .992
                     /codon_start=1
                     /product="regucalcin"
                     /protein_id="BAA07490.1"
                     /db_xref="GI:408807"
                     /translation="MSSIKIECVLRENYRCGESPVWEEASKCLLFVDIPSKTVCRWDS
                     ISNRVQRVGVDAPVSSVALRQSGGYVATIGTKFCALNWEDQSVFILAMVDEDKKNNRF
                     NDGKVDPAGRYFAGTMAEETAPAVLERHQGSLYSLFPDHSVKKYFNQVDISNGLDWSL
                     DHKIFYYIDSLSYTVDAFDYDLPTGQISNRRTVYKMEKDEQIPDGMCIDVEGKLWVAC
                     YNGGRVIRLDPETGKRLQTVKLPVDKTTSCCFGGKDYSEMYVTCARDGMSAEGLLRQP
                     DAGNIFKITGLGVKGIAPYSYAG"
     polyA_signal    1154. .1158
BASE COUNT      464 a    300 c    365 g    465 t
ORIGIN

  Query Match             9.6%;  Score 88.8;  DB 10;  Length 1594;
  Best Local Similarity   50.1%;  Pred. No. 1.4e-11;
  Matches  306;  Conservative    0;  Mismatches  293;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
```

```
Db        369 GTGGATGAAGATAAGAAAAACAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 428

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414

Db        429 TACTTTGCTGGTACCATGGCTGAGGAAACCGCCCCAGCTGTTCTGGAGCGGCACCAAGGG 488

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474

Db        489 TCCTTGTACTCCCTTTTTCCTGATCACAGTGTGAAGAAATACTTTAACCAAGTGGATATC 548

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534

Db        549 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 608

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594

Db        609 TACACTGTGGATGCCTTTGACTATGACCTGCCAACAGGACAGATTTCCAACCGCAGGACT 668

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654

Db        669 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGTT 722

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714

Db        723 GAGGGGAAGCTTTGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTAGATCCTGAG 782

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774

Db        783 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 839

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834

Db        840 TTTGGAGGGAAGGATTACTCTGAAATGTACGTGACATGTGCCAGGGATGGGATGAGCGCC 899

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894

Db        900 GAAGGTCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTTGGGGTC 959

Qy        895 AAAGGTTTCGC 905

Db        960 AAAGGAATTGC 970


RESULT 11
RNSMP30A
LOCUS       RNSMP30A                1594 bp    mRNA    linear   ROD 16-NOV-1993
DEFINITION  R.norvegicus SMP30 mRNA for senescence marker protein-30.
ACCESSION   X69021 S48114
VERSION     X69021.1  GI:57254
KEYWORDS    senescence marker protein-30.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Fujita,T., Shirasawa,T., Uchida,K. and Maruyama,N.
  TITLE     Isolation of cDNA clone encoding rat senescence marker protein-30
            (SMP30) and its tissue distribution
  JOURNAL   Biochim. Biophys. Acta 1132 (3), 297-305 (1992)
  MEDLINE   93041931
   PUBMED   1420310
REFERENCE   2  (bases 1 to 1594)
  AUTHORS   Fujita,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-1992) T. Fujita, Dept. o Molecular Biology, Tokyo
            Metropolitan Inst. of Gerontology, 35-2, Sakaecho, Itabashi-ku,
            Tokyo 173, JAPAN
FEATURES             Location/Qualifiers
     source          1. .1594
                     /organism="Rattus norvegicus"
                     /strain="Wistar/Sic"
                     /db_xref="taxon:10116"
                     /dev_stage="6 months"
     gene            88. .1588
                     /gene="SMP30"
     CDS             88. .987
                     /gene="SMP30"
                     /codon_start=1
                     /product="senescence marker  protein-30"
                     /protein_id="CAA48786.1"
                     /db_xref="GI:57255"
                     /db_xref="SWISS-PROT:Q03336"
                     /translation="MSSIKIECVLRENYRCGESPVWEEASKCLLFVDIPSKTVCRWDS
                     ISNRVQRVGVDAPVSSVALRQSGGYVATIGTKFCALNWEDQSVFILAMVDEDKKNNRF
                     NDGKVDPAGRYFAGTMAEETAPAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSL
                     DHKIFYYIDSLSYTVDAFDYDLPTGQISNRRTVYKMEKDEQIPDGMCIDVEGKLWVAC
                     YNGGRVIRLDPETGKRLQTVKLPVDKTTSCCFGGKDYSEMYVTCARDGMSAEGLLRQP
                     DAGNIFKITGLGVKGIAPYSYAG"
     polyA_signal    1148. .1153
                     /gene="SMP30"
     polyA_signal    1566. .1571
                     /gene="SMP30"
     polyA_site      1588
                     /gene="SMP30"
BASE COUNT      466 a    298 c    366 g    464 t
ORIGIN

  Query Match             9.3%;  Score 85.6;  DB 10;  Length 1594;
  Best Local Similarity   49.8%;  Pred. No. 9e-11;
  Matches  304;  Conservative    0;  Mismatches  295;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357

Db        364 GTGGATGAAGATAAGAAAAACAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 423

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414

Db        424 TACTTTGCTGGTACCATGGCTGAGGAAACGGCCCCAGCTGTTCTGGAGCGGCACCAAGGG 483

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474

Db        484 TCCTTGTACTCCCTTTTTCCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATC 543

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534

Db        544 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 603

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594

Db        604 TACACTGTGGATGCCTTTGACTATGACCTGCCAACAGGACAGATTTCCAACCGCAGAACT 663

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654

Db        664 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGTT 717

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714

Db        718 GAGGGGAAGCTTTGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTAGATCCTGAG 777

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774

Db        778 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 834

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834

Db        835 TTTGGAGGGAAGGATTACTCTGAAATGTACGTGACATGTGCCAGGGATGGGATGAGCGCC 894

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894

Db        895 GAAGGTCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTTGGGGTC 954

Qy        895 AAAGGTTTCGC 905

Db        955 AAAGGAATTGC 965
```

```
RESULT 12
AB037934
LOCUS       AB037934                1605 bp    mRNA    linear   ROD 09-FEB-2000
DEFINITION  Rattus norvegicus mRNA for regucalcin, complete cds.
ACCESSION   AB037934
VERSION     AB037934.1  GI:6970312
KEYWORDS    regucalcin.
SOURCE      Rattus norvegicus hepatoma cell_line:H4-II-E cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Misawa,H. and Yamaguchi,M.
  TITLE     The gene family encoding the calcium-binding protein regucalcin
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1605)
  AUTHORS   Yamaguchi,M. and Misawa,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-2000) Masayoshi Yamaguchi, Graduate School of
            Nutritional Sciences, University of Shizuoka, Laboratory of
            Endocrinology and Molecular Metabolism; 52-1 Yada, Shizuoka
            422-8526, Japan (E-mail:yamaguch@fns1.u-shizuoka-ken.ac.jp,
            Tel:+81-54-264-5580, Fax:+81-54-264-5580)
FEATURES             Location/Qualifiers
     source          1. .1605
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /cell_line="H4-II-E"
                     /cell_type="hepatoma"
     CDS             99. .998
                     /codon_start=1
                     /product="regucalcin"
                     /protein_id="BAA90692.1"
                     /db_xref="GI:6970313"
                     /translation="MSSIKIECVLRENYRCGESPVWEEASKCLLFVDIPSKTVCRWDS
                     ISNRVQRVGVDAPVSSVALRQSGGYVATIGTKFCALNWEDQSVFILAMVDEDKKNNRF
                     NDGKVDPAGRYFAGTMAEETAPAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSL
                     DHKIFYYIDSLSYTVDAFDYDLPTGQISNRRTVYKMEKDEQIPDGMCIDVEGKLWVAC
                     YNGGRVIRLDPETGKRLQTVKLPVDKTTSCCFGGKDYSEMYVTCARDGMSAEGLLRQP
                     DAGNIFKITGLGVKGIAPYSYAG"
BASE COUNT      471 a    300 c    367 g    467 t
ORIGIN

  Query Match             9.3%;  Score 85.6;  DB 10;  Length 1605;
  Best Local Similarity   49.8%;  Pred. No. 9e-11;
  Matches  304;  Conservative    0;  Mismatches  295;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||   | | |  ||||| ||  || |||||||||| || |  |||||   ||| |
Db        375 GTGGATGAAGATAAGAAAAACAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 434

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414
                |  | |||||  ||||||   ||  | |  |   | ||    |  | | |   ||
Db        435 TACTTTGCTGGTACCATGGCTGAGGAAACGGCCCCAGCTGTTCTGGAGCGGCACCAAGGG 494

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
              || ||      |     |||||  | |  || || |   ||      | |  | | | |||
Db        495 TCCTTGTACTCCCTTTTTCCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATC 554

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
              | |||||  | |  |||       ||    || |   | || || |||||       |
Db        555 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 614

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
                | || || |  |  || |||||      | |||  |   ||   |||  |   |  |
Db        615 TACACTGTGGATGCCTTTGACTATGACCTGCCAACAGGACAGATTTCCAACCGCAGAACT 674

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
              | |  |   | |||||  |||||        |   ||||||||      |||||||
Db        675 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGTT 728

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
              ||||| ||  | |||||  |||  || | | |||    || | |||    |   | | |
Db        729 GAGGGGAAGCTTTGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTAGATCCTGAG 788

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
                |  | ||     |||| | || || ||  | || | |  | |    || ||
Db        789 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 845

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
              ||||| ||   | |||    |||| || | || |||| ||| |   |  | | |   || |
Db        846 TTTGGAGGGAAGGATTACTCTGAAATGTACGTGACATGTGCCAGGGATGGGATGAGCGCC 905

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
                | |    |        | |  |||  || | | ||   |||||||| | || |||
Db        906 GAAGGTCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTTGGGGTC 965

Qy        895 AAAGGTTTCGC 905
              |||||  | ||
Db        966 AAAGGAATTGC 976


RESULT 13
D86217
LOCUS       D86217                  1109 bp    mRNA    linear   ROD 07-FEB-1999
DEFINITION  Mus musculus mRNA for regucalcin, complete cds.
ACCESSION   D86217
VERSION     D86217.1  GI:1483151
KEYWORDS    calcium-binding protein; regucalcin.
SOURCE      Mus musculus liver cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
  AUTHORS   Murata,T. and Yamaguchi,M.
  TITLE     Molecular cloning of the cDNA coding for regucalcin and its mRNA
            expression in mouse liver: the expression is stimulated by calcium
            administration
  JOURNAL   Mol. Cell. Biochem. 173 (1-2), 127-133 (1997)
  MEDLINE   97422495
REFERENCE   2  (bases 1 to 1109)
  AUTHORS   Yamaguchi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUN-1996) Masayoshi Yamaguchi, University of
            Shizuoka, Laboratory of Endocrinology and Molecular Metabolism,
            Graduate School of Nurtritional Sciences; 52-1 Yada, Shizuoka,
            Shizuoka 422, Japan (Tel:054-264-5580, Fax:054-264-5580)
COMMENT     Sequence updated (15-Jul-1996) by: Masayoshi Yamaguchi.
FEATURES             Location/Qualifiers
     source          1. .1109
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /tissue_type="liver"
     CDS             25. .924
                     /codon_start=1
                     /product="regucalcin"
                     /protein_id="BAA13046.1"
                     /db_xref="GI:1483152"
                     /translation="MSSIKVECVLRENYRCGESPVWEEASQSLLFVDIPSKIICRWDT
                     VSNQVQRVAVDAPVSSVALRQLGGYVATIGTKFCALNWENQSVFVLAMVDEDKKNNRF
                     NDGKVDPAGRYFAGTMAEETAPAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSL
                     DHKIFYYIDSLSYTVDAFDYDLQTGQISNRRIVYKMEKDEQIPDGMCIDAEGKLWVAC
                     YNGGRVIRLDPETGKRLQTVKLPVDKTTSCCFGGKDYSEMYVTCARDGLNAEGLLRQP
                     DAGNIFKITGLGVKGIAPYSYAG"
BASE COUNT      305 a    207 c    282 g    315 t
ORIGIN

  Query Match             8.7%;  Score 80.8;  DB 10;  Length 1109;
  Best Local Similarity   49.3%;  Pred. No. 1.5e-09;
  Matches  301;  Conservative    0;  Mismatches  298;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||   | | |  ||||||||  || ||||||||||| || |  |||||   ||| |
```

```
Db        301 GTGGATGAAGATAAGAAAAATAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 360

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414

Db        361 TACTTTGCTGGTACCATGGCTGAGGAAACGGCCCCAGCTGTTCTTGAGCGGCACCAAGGG 420

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474

Db        421 TCCTTGTACTCCCTCTTTCCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATC 480

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534

Db        481 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 540

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594

Db        541 TACACTGTGGATGCCTTTGACTATGACCTACAAACAGGACAGATTTCCAACCGCAGAATT 600

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654

Db        601 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGCT 654

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714

Db        655 GAGGGAAAGCTATGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTGGATCCTGAG 714

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774

Db        715 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 771

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834

Db        772 TTTGGAGGGAAAGATTACTCTGAAATGTATGTGACCTGTGCCAGGGATGGGTTGAATGCT 831

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894

Db        832 GAAGGCCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTCGGAGTC 891

Qy        895 AAAGGTTTCGC 905

Db        892 AAAGGAATTGC 902


RESULT 14
MMU28937
LOCUS       MMU28937               1573 bp    mRNA    linear   ROD 22-AUG-1996
DEFINITION  Mus musculus senescence marker protein-30 (SMP30) mRNA, complete
            cds.
ACCESSION   U28937
VERSION     U28937.1  GI:1143999
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1573)
  AUTHORS   Fujita,T., Shirasawa,T. and Maruyama,N.
  TITLE     Isolation and characterization of genomic and cDNA clones encoding
            mouse senescence marker protein-30 (SMP30)
  JOURNAL   Biochim. Biophys. Acta 1308 (1), 49-57 (1996)
  MEDLINE   96328264
   PUBMED   8765750
REFERENCE   2  (bases 1 to 1573)
  AUTHORS   Shirasawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1995) Takuji Shirasawa, Molecular Pathology,
            Tokyo Metropolitan Institute of Gerontology, 35-2, Sakaecho,
            Itabashi-ku, Tokyo 173, Japan
FEATURES             Location/Qualifiers
     source          1. .1573
                     /organism="Mus musculus"
                     /strain="C57Black/6"
                     /db_xref="taxon:10090"
                     /tissue_type="liver"
                     /dev_stage="adult"
     gene            1. .1573
                     /gene="SMP30"
     CDS             73. .972
                     /gene="SMP30"
                     /codon_start=1
                     /product="senescence marker protein-30"
                     /protein_id="AAC52721.1"
                     /db_xref="GI:1144000"
                     /translation="MSSIKVECVLRENYRCGESPVWEEASQSLLFVDIPSKIICRWDT
                     VSNQVQRVAVDAPVSSVALRQLGGYVATIGTKFCALNWENQSVFVLAMVDEDKKNNRF
                     NDGKVDPAGRYFAGTMAEETAPAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSL
                     DHKIFYYIDSLSYTVDAFDYDLQTGQISNRRIVYKMEKDEQIPDGMCIDAEGKLWVAC
                     YNGGRVIRLDPETGKRLQTVKLPVDKTTSCCFGGKDYSEMYVTCARDGLNAEGLLRQP
                     DAGNIFKITGLGVKGIAPYSYAG"
     polyA_signal    1132. .1137
                     /gene="SMP30"
     polyA_signal    1551. .1556
                     /gene="SMP30"
BASE COUNT      453 a    288 c    366 g    466 t
ORIGIN

  Query Match             8.7%;  Score 80.8;  DB 10;  Length 1573;
  Best Local Similarity   49.3%;  Pred. No. 1.4e-09;
  Matches  301;  Conservative    0;  Mismatches  298;  Indels   12;  Gaps    3;

Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357

Db        349 GTGGATGAAGATAAGAAAAATAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 408

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414

Db        409 TACTTTGCTGGTACCATGGCTGAGGAAACGGCCCCAGCTGTTCTTGAGCGGCACCAAGGG 468

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474

Db        469 TCCTTGTACTCCCTCTTTCCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATC 528

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534

Db        529 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 588

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594

Db        589 TACACTGTGGATGCCTTTGACTATGACCTACAAACAGGACAGATTTCCAACCGCAGAATT 648

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654

Db        649 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGCT 702

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714

Db        703 GAGGGAAAGCTATGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTGGATCCTGAG 762

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774

Db        763 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 819

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834

Db        820 TTTGGAGGGAAAGATTACTCTGAAATGTATGTGACCTGTGCCAGGGATGGGTTGAATGCT 879

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894

Db        880 GAAGGCCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTCGGAGTC 939

Qy        895 AAAGGTTTCGC 905

Db        940 AAAGGAATTGC 950


RESULT 15
```

```
BC012710
LOCUS       BC012710                1598 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to regucalcin, clone MGC:14006 IMAGE:4210374,
            mRNA, complete cds.
ACCESSION   BC012710
VERSION     BC012710.1  GI:15215230
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1598)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 18 Row: n Column: 22
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6677738.
FEATURES             Location/Qualifiers
     source          1. .1598
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:14006 IMAGE:4210374"
                     /tissue_type="Liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             86. .985
                     /codon_start=1
                     /product="Similar to regucalcin"
                     /protein_id="AAH12710.1"
                     /db_xref="GI:15215231"
                     /translation="MSSIKVECVLRENYRCGESPVWEEASQSLLFVDIPSKIICRWDT
                     VSNQVQRVAVDAPVSSVALRQLGGYVATIGTKFCALNWENQSVFVLAMVDEDKKNNRF
                     NDGKVDPAGRYFAGTMAEETAPAVLERHQGSLYSLFPDHSVKKYFDQVDISNGLDWSL
                     DHKIFYYIDSLSYTVDAFDYDLQTGQISNRRIVYKMEKDEQIPDGMCIDAEGKLWVAC
                     YNGGRVIRLDPETGKRLQTVKLPVDKTTSCCFGGKDYSEMYVTCARDGLNAEGLLRQP
                     DAGNIFKITGLGVKGIAPYSYAG"
BASE COUNT      466 a    293 c    370 g    469 t
ORIGIN
```

Query Match 8.7%; Score 80.8; DB 10; Length 1598;
Best Local Similarity 49.3%; Pred. No. 1.4e-09;
Matches 301; Conservative 0; Mismatches 298; Indels 12; Gaps 3;

```
Qy        298 GTCGAACCACACATTAAAAATAACAGACTCAATGATGGCAAAGCAGATCCCCTTGGCAAT 357
              || ||   | | |  |||||||  || ||||||||| || |  |||||   ||| |
Db        362 GTGGATGAAGATAAGAAAAATAATCGATTCAATGATGGGAAGGTGGATCCTGCTGGGAGA 421

Qy        358 CTATGGACAGGTACAATGGCTATTGACGCTGGTCTCCCCGTA---GGACCGGTCACTGGC 414
                |   | ||||| ||||||   ||  | |  |   | ||    |  | |  |   ||
Db        422 TACTTTGCTGGTACCATGGCTGAGGAAACGGCCCCAGCTGTTCTTGAGCGGCACCAAGGG 481

Qy        415 AGTTTATATCATTTAGGGGCTGATAAAAAGGTAAAAATGCACGAGAGCAACATAGCTATA 474
                || ||     |     ||||| | |  || || |   ||     | |  | | |||

Db        482 TCCTTGTACTCCCTCTTTCCTGATCACAGTGTGAAGAAATACTTTGACCAAGTGGATATC 541

Qy        475 GCAAATGGGCTCGCGTGGAGTAATGATTTGAAGAAAATGTATTATATTGATTCGGGGAAA 534
              | |||||  | |  |||      ||    || |   | || || |||||    |
Db        542 TCCAATGGTTTGGATTGGTCCCTGGACCATAAAATCTTCTACTACATTGACAGCCTGTCC 601

Qy        535 AGAAGAGTAGACGAGTACGATTATGATGCTTCTACATTATCCATCAGCAATCAACGGCCA 594
                |  || || |  |  || ||||||       |||  |   ||   |||  |   |
Db        602 TACACTGTGGATGCCTTTGACTATGACCTACAAACAGGACAGATTTCCAACCGCAGAATT 661

Qy        595 TTATTTACTTTTGAAAAGCATGAAGTGCCTGGATATCCAGATGGTCAAACAATTGATGAG 654
              | |  | |  | |||||  |||||        |   |||||||        |||||||
Db        662 GTTTACAAGATGGAAAAAGATGAA------CAAATCCCAGATGGAATGTGCATTGATGCT 715

Qy        655 GAGGGTAATTTATGGGTTGCCGTTTTCCAAGGACAGCGAATTATTAAAATCAGTACCCAA 714
              ||||| ||  |||||||  |||  || | | |||    || | |||    |  | |  |
Db        716 GAGGGAAAGCTATGGGTGGCCTGTTACAATGGAGGAAGAGTAATTCGCCTGGATCCTGAG 775

Qy        715 CAACCGGAAGTGTTACTGGATACCGTAAAAATACCAGATCCTCAGGTCACCTCTGTAGCA 774
                | | ||     |||| | || || ||  | || | |  | |    || ||
Db        776 ACAGGGAAA---AGACTGCAAACTGTGAAGTTGCCTGTTGATAAAACAACTTCATGCTGC 832

Qy        775 TTTGGCGGTCCGAATTTGGATGAACTGCATGTAACATCTGCTGGTCTTCAGCTTGACGAC 834
              ||||| ||     |||    |||| || |||| || | |||  |   |  | | | | |
Db        833 TTTGGAGGGAAAGATTACTCTGAAATGTATGTGACCTGTGCCAGGGATGGGTTGAATGCT 892

Qy        835 AGTTCTTTNGACAAAAGTTTAGTTAATGGGCACGTCTACAGAGTAACAGGTTTAGGCGTC 894
                    |   |          | |  |||  || | | ||   |||||||| | || |||
Db        893 GAAGGCCTTTTGAGGCAGCCTGATGCTGGTAACATTTTCAAGATAACAGGTCTCGGAGTC 952

Qy        895 AAAGGTTTCGC 905
              |||||  | ||
Db        953 AAAGGAATTGC 963
```

Search completed: August 1, 2003, 12:51:16
Job time : 2658 secs